BUNDESREPUBLIK DEUTSCHLAND

CERTIFIED COPY OF PRIORITY DOCUMENT

Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

Aktenzeichen:	199 30 429.7
Anmeldetag:	01. Juli 1999
Anmelder/Inhaber:	BASF Aktiengesellschaft, Ludwigshafen/DE
Bezeichnung:	Corynebacterium glutamicum genes encoding regulatory proteins
IPC:	C 07 K, C 12 N

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Anmeldung.

München, den 13. Juli 2000
Deutsches Patent- und Markenamt
Der Präsident
Im Auftrag

Weihmayr

CORYNEBACTERIUM GLUTAMICUM GENES ENCODING REGULATORY PROTEINS

Background of the Invention

Certain products and by-products of naturally-occurring metabolic processes in cells have utility in a wide array of industries, including the food, feed, cosmetics, and pharmaceutical industries. These molecules, collectively termed 'fine chemicals', include organic acids, both proteinogenic and non-proteinogenic amino acids, nucleotides and nucleosides, lipids and fatty acids, diols, carbohydrates, aromatic compounds, vitamins and cofactors, and enzymes. Their production is most conveniently performed through the large-scale culture of bacteria developed to produce and secrete large quantities of one or more desired molecules. One particularly useful organism for this purpose is *Corynebacterium glutamicum*, a gram positive, nonpathogenic bacterium. Through strain selection, a number of mutant strains have been developed which produce an array of desirable compounds. However, selection of strains improved for the production of a particular molecule is a time-consuming and difficult process.

Summary of the Invention

This invention provides novel nucleic acid molecules which may be used to identify or classify *Corynebacterium glutamicum* or related species of bacteria. *C. glutamicum* is a gram positive, aerobic bacterium which is commonly used in industry for the large-scale production of a variety of fine chemicals, and also for the degradation of hydrocarbons (such as in petroleum spills) and for the oxidation of terpenoids. The nucleic acid molecules therefore can be used to identify microorganisms which can be used to produce fine chemicals, e.g., by fermentation processes. While *C. glutamicum* itself is nonpathogenic, it is related to other Corynebacterium species, such as *Corynebacterium diphtheriae* (the causative agent of diphtheria), which are important human pathogens. The ability to identify the presence of Corynebacterium species therefore also can have significant clinical relevance, e.g., diagnostic applications. Further, these nucleic acid molecules may serve as reference points for the mapping of the *C. glutamicum* genome, or of genomes of related organisms.

These novel nucleic acid molecules encode proteins, referred to herein as metabolic regulatory (MR) proteins. These MR proteins are capable of, for example, performing a function involved in the transcriptional, translational, or posttranslational regulation of proteins important for the normal metabolic functioning of cells. Given the availability of cloning vectors for use in *Corynebacterium glutamicum*, such as those

disclosed in Sinskey et al., U.S. Patent No. 4,649,119, and techniques for genetic manipulation of *C. glutamicum* and the related *Brevibacterium* species (e.g., *lactofermentum*) (Yoshihama et al., *J Bacteriol.* 162: 591-597 (1985); Katsumata et al., *J. Bacteriol.* 159: 306-311 (1984); and Santamaria et al., *J. Gen. Microbiol.* 130: 2237-2246 (1984)), the nucleic acid molecules of the invention may be utilized in the genetic engineering of this organism to make it a better or more efficient producer of one or more fine chemicals.

This improved yield, production and/or efficiency of production of a fine chemical may be due to a direct effect of manipulation of a gene of the invention, or it may be due to an indirect effect of such manipulation. Specifically, alterations in *C. glutamicum* MR proteins which normally regulate the yield, production and/or efficiency of production of a fine chemical metabolic pathways may have a direct impact on the overall production or rate of production of one or more of these desired compounds from this organism. Alterations in the proteins involved in these metabolic pathways may also have an indirect impact on the yield, production and/or efficiency of production of a desired fine chemical. Regulation of metabolism is necessarily complex, and the regulatory mechanisms governing different pathways may intersect at multiple points such that more than one pathway can be rapidly adjusted in accordance with a particular cellular event. This enables the modification of a regulatory protein for one pathway to have an impact on the regulation of many other pathways as well, some of which may be involved in the biosynthesis or degradation of a desired fine chemical. In this indirect fashion, the modulation of action of an MR protein may have an impact on the production of a fine chemical produced by a pathway different from one which that MR protein directly regulates.

The nucleic acid and protein molecules of the invention may be utilized to directly improve the yield, production, and/or efficiency of production of one or more desired fine chemicals from *Corynebacterium glutamicum*. Using recombinant genetic techniques well known in the art, one or more of the regulatory proteins of the invention may be manipulated such that its function is modulated. For example, the mutation of an MR protein involved in the repression of transcription of a gene encoding an enzyme which is required for the biosynthesis of an amino acid such that it no longer is able to repress transcription may result in an increase in production of that amino acid. Similarly, the alteration of activity of an MR protein resulting in increased translation or activating posttranslational modification of a *C. glutamicum* protein involved in the biosynthesis of a desired fine chemical may in turn increase the production of that chemical. The opposite situation may also be of benefit: by increasing the repression of transcription or translation, or by posttranslational negative modification of a *C.*

glutamicum protein involved in the regulation of a degradative pathway for a compound, one may increase the production of this chemical. In each case, the overall yield or rate of production of the desired fine chemical may be increased.

It is also possible that such alterations in the protein and nucleotide molecules of the invention may improve the yield, production, and/or efficiency of production of fine chemicals through indirect mechanisms. The metabolism of any one compound is necessarily intertwined with other biosynthetic and degradative pathways within the cell, and necessary cofactors, intermediates, or substrates in one pathway are likely supplied or limited by another such pathway. Therefore, by modulating the activity of one or more of the regulatory proteins of the invention, the production or efficiency of activity of another fine chemical biosynthetic or degradative pathway may be impacted. Further, the manipulation of one or more regulatory proteins may increase the overall ability of the cell to grow and multiply in culture, particularly in large-scale fermentative culture, where growth conditions may be suboptimal. For example, by mutating an MR protein of the invention which would normally cause a repression in the biosynthesis of nucleotides in response to suboptimal extracellular supplies of nutrients (thereby preventing cell division) such that it is decreased in repressor ability, one may increase the biosynthesis of nucleotides and perhaps increase cell division. Changes in MR proteins which result in increased cell growth and division in culture may result in an increase in yield, production, and/or efficiency of production of one or more desired fine chemicals from the culture, due at least to the increased number of cells producing the chemical in the culture.

The invention provides novel nucleic acid molecules which encode proteins, referred to herein as metabolic pathway proteins (MR), which are capable of, for example, performing an enzymatic step involved in the transcriptional, translational, or posttranslational regulation of metabolic pathways in *C. glutamicum*. Nucleic acid molecules encoding an MR protein are referred to herein as MR nucleic acid molecules. In a preferred embodiment, the MR protein participates in the transcriptional, translational, or posttranslational regulation of one or more metabolic pathways. Examples of such proteins include those encoded by the genes set forth in Table 1.

Accordingly, one aspect of the invention pertains to isolated nucleic acid molecules (e.g., cDNAs) comprising a nucleotide sequence encoding an MR protein or biologically active portions thereof, as well as nucleic acid fragments suitable as primers or hybridization probes for the detection or amplification of MR-encoding nucleic acid (e.g., DNA or mRNA). In particularly preferred embodiments, the isolated nucleic acid molecule comprises one of the nucleotide sequences set forth in Appendix A or the coding region or a complement thereof of one of these nucleotide sequences. In other

particularly preferred embodiments, the isolated nucleic acid molecule of the invention comprises a nucleotide sequence which hybridizes to or is at least about 50%, preferably at least about 60%, more preferably at least about 70%, 80% or 90%, and even more preferably at least about 95%, 96%, 97%, 98%, 99% or more homologous to a nucleotide sequence set forth in Appendix A, or a portion thereof. In other preferred embodiments, the isolated nucleic acid molecule encodes one of the amino acid sequences set forth in Appendix B. The preferred MR proteins of the present invention also preferably possess at least one of the MR activities described herein.

In another embodiment, the isolated nucleic acid molecule encodes a protein or portion thereof wherein the protein or portion thereof includes an amino acid sequence which is sufficiently homologous to an amino acid sequence of Appendix B, e.g., sufficiently homologous to an amino acid sequence of Appendix B such that the protein or portion thereof maintains an MR activity. Preferably, the protein or portion thereof encoded by the nucleic acid molecule maintains the ability to transcriptionally, translationally, or posttranslationally regulate a metabolic pathway in *C. glutamicum*. In one embodiment, the protein encoded by the nucleic acid molecule is at least about 50%, preferably at least about 60%, and more preferably at least about 70%, 80%, or 90% and most preferably at least about 95%, 96%, 97%, 98%, or 99% or more homologous to an amino acid sequence of Appendix B (e.g., an entire amino acid sequence selected from those sequences set forth in Appendix B). In another preferred embodiment, the protein is a full length *C. glutamicum* protein which is substantially homologous to an entire amino acid sequence of Appendix B (encoded by an open reading frame shown in Appendix A).

In another preferred embodiment, the isolated nucleic acid molecule is derived from *C. glutamicum* and encodes a protein (e.g., an MR fusion protein) which includes a biologically active domain which is at least about 50% or more homologous to one of the amino acid sequences of Appendix B and is able to transcriptionally, translationally, or posttranslationally regulate a metabolic pathway in *C. glutamicum*, or has one or more of the activities set forth in Table 1, and which also includes heterologous nucleic acid sequences encoding a heterologous polypeptide or regulatory regions.

In another embodiment, the isolated nucleic acid molecule is at least 15 nucleotides in length and hybridizes under stringent conditions to a nucleic acid molecule comprising a nucleotide sequence of Appendix A. Preferably, the isolated nucleic acid molecule corresponds to a naturally-occurring nucleic acid molecule. More preferably, the isolated nucleic acid encodes a naturally-occurring *C. glutamicum* MR protein, or a biologically active portion thereof.

Another aspect of the invention pertains to vectors, e.g., recombinant expression vectors, containing the nucleic acid molecules of the invention, and host cells into which such vectors have been introduced. In one embodiment, such a host cell is used to produce an MR protein by culturing the host cell in a suitable medium. The MR protein can be then isolated from the medium or the host cell.

Yet another aspect of the invention pertains to a genetically altered microorganism in which an MR gene has been introduced or altered. In one embodiment, the genome of the microorganism has been altered by introduction of a nucleic acid molecule of the invention encoding wild-type or mutated MR sequence as a transgene. In another embodiment, an endogenous MR gene within the genome of the microorganism has been altered, e.g., functionally disrupted, by homologous recombination with an altered MR gene. In a preferred embodiment, the microorganism belongs to the genus *Corynebacterium* or *Brevibacterium*, with *Corynebacterium glutamicum* being particularly preferred. In a preferred embodiment, the microorganism is also utilized for the production of a desired compound, such as an amino acid, with lysine being particularly preferred.

Still another aspect of the invention pertains to an isolated MR protein or a portion, e.g., a biologically active portion, thereof. In a preferred embodiment, the isolated MR protein or portion thereof transcriptionally, translationally, or posttranslationally regulates one or more metabolic pathways in *C. glutamicum*. In another preferred embodiment, the isolated MR protein or portion thereof is sufficiently homologous to an amino acid sequence of Appendix B such that the protein or portion thereof maintains the ability to transcriptionally, translationally, or posttranslationally regulate one or more metabolic pathways in *C glutamicum*.

The invention also provides an isolated preparation of an MR protein. In preferred embodiments, the MR protein comprises an amino acid sequence of Appendix B. In another preferred embodiment, the invention pertains to an isolated full length protein which is substantially homologous to an entire amino acid sequence of Appendix B (encoded by an open reading frame set forth in Appendix A). In yet another embodiment, the protein is at least about 50%, preferably at least about 60%, and more preferably at least about 70%, 80%, or 90%, and most preferably at least about 95%, 96%, 97%, 98%, or 99% or more homologous to an entire amino acid sequence of Appendix B. In other embodiments, the isolated MR protein comprises an amino acid sequence which is at least about 50% or more homologous to one of the amino acid sequences of Appendix B and is able to transcriptionally, translationally, or posttranslationally regulate one or more metabolic pathways in *C. glutamicum*, or has one or more of the activities set forth in Table 1.

Alternatively, the isolated MR protein can comprise an amino acid sequence which is encoded by a nucleotide sequence which hybridizes, e.g., hybridizes under stringent conditions, or is at least about 50%, preferably at least about 60%, more preferably at least about 70%, 80%, or 90%, and even more preferably at least about 95%, 96%, 97%, 98,%, or 99% or more homologous, to a nucleotide sequence of Appendix B. It is also preferred that the preferred forms of MR proteins also have one or more of the MR bioactivities described herein.

The MR polypeptide, or a biologically active portion thereof, can be operatively linked to a non-MR polypeptide to form a fusion protein. In preferred embodiments, this fusion protein has an activity which differs from that of the MR protein alone. In other preferred embodiments, this fusion protein transcriptionally, translationally, or posttranslationally regulates one or more metabolic pathways in *C. glutamicum*. In particularly preferred embodiments, integration of this fusion protein into a host cell modulates production of a desired compound from the cell.

Another aspect of the invention pertains to a method for producing a fine chemical. This method involves the culturing of a cell containing a vector directing the expression of an MR nucleic acid molecule of the invention, such that a fine chemical is produced. In a preferred embodiment, this method further includes the step of obtaining a cell containing such a vector, in which a cell is transfected with a vector directing the expression of an MR nucleic acid. In another preferred embodiment, this method further includes the step of recovering the fine chemical from the culture. In a particularly preferred embodiment, the cell is from the genus *Corynebacterium* or *Brevibacterium*, or is selected from those strains set forth in Table 3.

Another aspect of the invention pertains to methods for modulating production of a molecule from a microorganism. Such methods include contacting the cell with an agent which modulates MR protein activity or MR nucleic acid expression such that a cell associated activity is altered relative to this same activity in the absence of the agent. In a preferred embodiment, the cell is modulated for one or more *C. glutamicum* metabolic pathway regulatory systems, such that the yields or rate of production of a desired fine chemical by this microorganism is improved. The agent which modulates MR protein activity can be an agent which stimulates MR protein activity or MR nucleic acid expression. Examples of agents which stimulate MR protein activity or MR nucleic acid expression include small molecules, active MR proteins, and nucleic acids encoding MR proteins that have been introduced into the cell. Examples of agents which inhibit MR activity or expression include small molecules and antisense MR nucleic acid molecules.

Another aspect of the invention pertains to methods for modulating yields of a desired compound from a cell, involving the introduction of a wild-type or mutant MR gene into a cell, either maintained on a separate plasmid or integrated into the genome of the host cell. If integrated into the genome, such integration can be random, or it can take place by homologous recombination such that the native gene is replaced by the introduced copy, causing the production of the desired compound from the cell to be modulated. In a preferred embodiment, said yields are increased. In another preferred embodiment, said chemical is a fine chemical. In a particularly preferred embodiment, said fine chemical is an amino acid. In especially preferred embodiments, said amino acid is L-lysine.

Detailed Description of the Invention

The present invention provides MR nucleic acid and protein molecules which are involved in the regulation of metabolism in *Corynebacterium glutamicum*, including regulation of fine chemical metabolism. The molecules of the invention may be utilized in the modulation of production of fine chemicals from microorganisms, such as *C glutamicum*, either directly (e.g., where modulation of the activity of a lysine biosynthesis regulatory protein has a direct impact on the yield, production, and/or efficiency of production of lysine from that organism), or an indirect impact which nonetheless results in an increase in yield, production, and/or efficiency of production of the desired compound (e.g., where modulation of the regulation of a nucleotide biosynthesis protein has an impact on the production of an organic acid or a fatty acid from the bacterium, perhaps due to concomitant regulatory alterations in the biosynthetic or degradation pathways for these chemicals in response to the altered regulation of nucleotide biosynthesis). Aspects of the invention are further explicated below.

I. Fine Chemicals

The term 'fine chemical' is art-recognized and includes molecules produced by an organism which have applications in various industries, such as, but not limited to, the pharmaceutical, agriculture, and cosmetics industries. Such compounds include organic acids, such as tartaric acid, itaconic acid, and diaminopimelic acid, both proteinogenic and non-proteinogenic amino acids, purine and pyrimidine bases, nucleosides, and nucleotides (as described e.g. in Kuninaka, A. (1996) Nucleotides and related compounds, p. 561-612, in Biotechnology vol. 6, Rehm et al., eds. VCH: Weinheim, and references contained therein), lipids, both saturated and unsaturated fatty acids (e.g., arachidonic acid), diols (e.g., propane diol, and butane diol), carbohydrates

(e.g., hyaluronic acid and trehalose), aromatic compounds (e.g., aromatic amines, vanillin, and indigo), vitamins and cofactors (as described in Ullmann's Encyclopedia of Industrial Chemistry, vol. A27, "Vitamins", p. 443-613 (1996) VCH: Weinheim and references therein; and Ong, A.S., Niki, E. & Packer, L. (1995) "Nutrition, Lipids, Health, and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia, and the Society for Free Radical Research – Asia, held Sept. 1-3, 1994 at Penang, Malaysia, AOCS Press, (1995)), enzymes, and all other chemicals described in Gutcho (1983) Chemicals by Fermentation, Noyes Data Corporation, ISBN: 0818805086 and references therein. The metabolism and uses of certain of these fine chemicals are further explicated below.

A. *Amino Acid Metabolism and Uses*

Amino acids comprise the basic structural units of all proteins, and as such are essential for normal cellular functioning in all organisms. The term "amino acid" is art-recognized. The proteinogenic amino acids, of which there are 20 species, serve as structural units for proteins, in which they are linked by peptide bonds, while the nonproteinogenic amino acids (hundreds of which are known) are not normally found in proteins (see Ulmann's Encyclopedia of Industrial Chemistry, vol. A2, p. 57-97 VCH: Weinheim (1985)). Amino acids may be in the D- or L- optical configuration, though L-amino acids are generally the only type found in naturally-occurring proteins. Biosynthetic and degradative pathways of each of the 20 proteinogenic amino acids have been well characterized in both prokaryotic and eukaryotic cells (see, for example, Stryer, L. Biochemistry, 3rd edition, pages 578-590 (1988)). The 'essential' amino acids (histidine, isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan, and valine), so named because they are generally a nutritional requirement due to the complexity of their biosynthesis, are readily converted by simple biosynthetic pathways to the remaining 11 'nonessential' amino acids (alanine, arginine, asparagine, aspartate, cysteine, glutamate, glutamine, glycine, proline, serine, and tyrosine). Higher animals do retain the ability to synthesize some of these amino acids, but the essential amino acids must be supplied from the diet in order for normal protein synthesis to occur.

Aside from their function in protein biosynthesis, these amino acids are interesting chemicals in their own right, and many have been found to have various applications in the food, feed, chemical, cosmetics, agriculture, and pharmaceutical industries. Lysine is an important amino acid in the nutrition not only of humans, but also of monogastric animals such as poultry and swine. Glutamate is most commonly used as a flavor additive (mono-sodium glutamate, MSG) and is widely used throughout the food industry, as are aspartate, phenylalanine, glycine, and cysteine. Glycine, L-

methionine and tryptophan are all utilized in the pharmaceutical industry. Glutamine, valine, leucine, isoleucine, histidine, arginine, proline, serine and alanine are of use in both the pharmaceutical and cosmetics industries. Threonine, tryptophan, and D/ L-methionine are common feed additives. (Leuchtenberger, W. (1996) Amino aids – technical production and use, p. 466-502 in Rehm et al. (eds.) Biotechnology vol. 6, chapter 14a, VCH: Weinheim). Additionally, these amino acids have been found to be useful as precursors for the synthesis of synthetic amino acids and proteins, such as N-acetylcysteine, S-carboxymethyl-L-cysteine, (S)-5-hydroxytryptophan, and others described in Ulmann's Encyclopedia of Industrial Chemistry, vol. A2, p. 57-97, VCH: Weinheim, 1985.

The biosynthesis of these natural amino acids in organisms capable of producing them, such as bacteria, has been well characterized (for review of bacterial amino acid biosynthesis and regulation thereof, see Umbarger, H.E.(1978) *Ann. Rev. Biochem.* 47: 533-606). Glutamate is synthesized by the reductive amination of α-ketoglutarate, an intermediate in the citric acid cycle. Glutamine, proline, and arginine are each subsequently produced from glutamate. The biosynthesis of serine is a three-step process beginning with 3-phosphoglycerate (an intermediate in glycolysis), and resulting in this amino acid after oxidation, transamination, and hydrolysis steps. Both cysteine and glycine are produced from serine; the former by the condensation of homocysteine with serine, and the latter by the transferal of the side-chain β-carbon atom to tetrahydrofolate, in a reaction catalyzed by serine transhydroxymethylase. Phenylalanine, and tyrosine are synthesized from the glycolytic and pentose phosphate pathway precursors erythrose 4-phosphate and phosphoenolpyruvate in a 9-step biosynthetic pathway that differ only at the final two steps after synthesis of prephenate. Tryptophan is also produced from these two initial molecules, but its synthesis is an 11-step pathway. Tyrosine may also be synthesized from phenylalanine, in a reaction catalyzed by phenylalanine hydroxylase. Alanine, valine, and leucine are all biosynthetic products of pyruvate, the final product of glycolysis. Aspartate is formed from oxaloacetate, an intermediate of the citric acid cycle. Asparagine, methionine, threonine, and lysine are each produced by the conversion of aspartate. Isoleucine is formed from threonine. A complex 9-step pathway results in the production of histidine from 5-phosphoribosyl-1-pyrophosphate, an activated sugar.

Amino acids in excess of the protein synthesis needs of the cell cannot be stored, and are instead degraded to provide intermediates for the major metabolic pathways of the cell (for review see Stryer, L. Biochemistry 3rd ed. Ch. 21 "Amino Acid Degradation and the Urea Cycle" p. 495-516 (1988)). Although the cell is able to convert unwanted amino acids into useful metabolic intermediates, amino acid production is costly in

terms of energy, precursor molecules, and the enzymes necessary to synthesize them. Thus it is not surprising that amino acid biosynthesis is regulated by feedback inhibition, in which the presence of a particular amino acid serves to slow or entirely stop its own production (for overview of feedback mechanisms in amino acid biosynthetic pathways, see Stryer, L. Biochemistry, 3rd ed. Ch. 24: "Biosynthesis of Amino Acids and Heme" p. 575-600 (1988)). Thus, the output of any particular amino acid is limited by the amount of that amino acid present in the cell.

B. Vitamin, Cofactor, and Nutraceutical Metabolism and Uses

Vitamins, cofactors, and nutraceuticals comprise another group of molecules which the higher animals have lost the ability to synthesize and so must ingest, although they are readily synthesized by other organisms such as bacteria. These molecules are either bioactive substances themselves, or are precursors of biologically active substances which may serve as electron carriers or intermediates in a variety of metabolic pathways. Aside from their nutritive value, these compounds also have significant industrial value as coloring agents, antioxidants, and catalysts or other processing aids. (For an overview of the structure, activity, and industrial applications of these compounds, see, for example, Ullman's Encyclopedia of Industrial Chemistry, "Vitamins" vol. A27, p. 443-613, VCH: Weinheim, 1996.) The term "vitamin" is art-recognized, and includes nutrients which are required by an organism for normal functioning, but which that organism cannot synthesize by itself. The group of vitamins may encompass cofactors and nutraceutical compounds. The language "cofactor" includes nonproteinaceous compounds required for a normal enzymatic activity to occur. Such compounds may be organic or inorganic; the cofactor molecules of the invention are preferably organic. The term "nutraceutical" includes dietary supplements having health benefits in plants and animals, particularly humans. Examples of such molecules are vitamins, antioxidants, and also certain lipids (e.g., polyunsaturated fatty acids).

The biosynthesis of these molecules in organisms capable of producing them, such as bacteria, has been largely characterized (Ullman's Encyclopedia of Industrial Chemistry, "Vitamins" vol. A27, p. 443-613, VCH: Weinheim, 1996; Michal, G. (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley & Sons; Ong, A.S., Niki, E. & Packer, L. (1995) "Nutrition, Lipids, Health, and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia, and the Society for Free Radical Research – Asia, held Sept. 1-3, 1994 at Penang, Malaysia, AOCS Press: Champaign, IL X, 374 S).

Thiamin (vitamin B_1) is produced by the chemical coupling of pyrimidine and thiazole moieties. Riboflavin (vitamin B_2) is synthesized from guanosine-5'-triphosphate (GTP) and ribose-5'-phosphate. Riboflavin, in turn, is utilized for the synthesis of flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD). The family of compounds collectively termed 'vitamin B_6' (e.g., pyridoxine, pyridoxamine, pyridoxa-5'-phosphate, and the commercially used pyridoxin hydrochloride) are all derivatives of the common structural unit, 5-hydroxy-6-methylpyridine. Pantothenate (pantothenic acid, (R)-(+)-N-(2.4-dihydroxy-3.3-dimethyl-1-oxobutyl)-β-alanine) can be produced either by chemical synthesis or by fermentation. The final steps in pantothenate biosynthesis consist of the ATP-driven condensation of β-alanine and pantoic acid. The enzymes responsible for the biosynthesis steps for the conversion to pantoic acid, to β-alanine and for the condensation to panthotenic acid are known. The metabolically active form of pantothenate is Coenzyme A, for which the biosynthesis proceeds in 5 enzymatic steps. Pantothenate, pyridoxal-5'-phosphate, cysteine and ATP are the precursors of Coenzyme A. These enzymes not only catalyze the formation of panthothante, but also the production of (R)-pantoic acid, (R)-pantolacton, (R)-panthenol (provitamin B_5), pantetheine (and its derivatives) and coenzyme A.

Biotin biosynthesis from the precursor molecule pimeloyl-CoA in microorganisms has been studied in detail and several of the genes involved have been identified. Many of the corresponding proteins have been found to also be involved in Fe-cluster synthesis and are members of the nifS class of proteins. Lipoic acid is derived from octanoic acid, and serves as a coenzyme in energy metabolism, where it becomes part of the pyruvate dehydrogenase complex and the α-ketoglutarate dehydrogenase complex. The folates are a group of substances which are all derivatives of folic acid, which is turn is derived from L-glutamic acid, p-amino-benzoic acid and 6-methylpterin. The biosynthesis of folic acid and its derivatives, starting from the metabolism intermediates guanosine-5'-triphosphate (GTP), L-glutamic acid and p-amino-benzoic acid has been studied in detail in certain microorganisms.

Corrinoids (such as the cobalamines and particularly vitamin B_{12}) and porphyrines belong to a group of chemicals characterized by a tetrapyrole ring system. The biosynthesis of vitamin B_{12} is sufficiently complex that it has not yet been completely characterized, but many of the enzymes and substrates involved are now known. Nicotinic acid (nicotinate), and nicotinamide are pyridine derivatives which are also termed 'niacin'. Niacin is the precursor of the important coenzymes NAD (nicotinamide adenine dinucleotide) and NADP (nicotinamide adenine dinucleotide phosphate) and their reduced forms.

The large-scale production of these compounds has largely relied on cell-free chemical syntheses, though some of these chemicals have also been produced by large-scale culture of microorganisms, such as riboflavin, Vitamin B_6, pantothenate, and biotin. Only Vitamin B_{12} is produced solely by fermentation, due to the complexity of its synthesis. *In vitro* methodologies require significant inputs of materials and time, often at great cost.

C. Purine, Pyrimidine, Nucleoside and Nucleotide Metabolism and Uses

Purine and pyrimidine metabolism genes and their corresponding proteins are important targets for the therapy of tumor diseases and viral infections. The language "purine" or "pyrimidine" includes the nitrogenous bases which are constituents of nucleic acids, co-enzymes, and nucleotides. The term "nucleotide" includes the basic structural units of nucleic acid molecules, which are comprised of a nitrogenous base, a pentose sugar (in the case of RNA, the sugar is ribose; in the case of DNA, the sugar is D-deoxyribose), and phosphoric acid. The language "nucleoside" includes molecules which serve as precursors to nucleotides, but which are lacking the phosphoric acid moiety that nucleotides possess. By inhibiting the biosynthesis of these molecules, or their mobilization to form nucleic acid molecules, it is possible to inhibit RNA and DNA synthesis; by inhibiting this activity in a fashion targeted to cancerous cells, the ability of tumor cells to divide and replicate may be inhibited. Additionally, there are nucleotides which do not form nucleic acid molecules, but rather serve as energy stores (i.e., AMP) or as coenzymes (i.e., FAD and NAD).

Several publications have described the use of these chemicals for these medical indications, by influencing purine and/or pyrimidine metabolism (e.g. Christopherson, R.I. and Lyons, S.D. (1990) "Potent inhibitors of *de novo* pyrimidine and purine biosynthesis as chemotherapeutic agents." *Med. Res. Reviews* 10: 505-548). Studies of enzymes involved in purine and pyrimidine metabolism have been focused on the development of new drugs which can be used, for example, as immunosuppressants or anti-proliferants (Smith, J.L., (1995) "Enzymes in nucleotide synthesis." *Curr Opin. Struct. Biol.* 5: 752-757; (1995) *Biochem Soc. Transact.* 23: 877-902). However, purine and pyrimidine bases, nucleosides and nucleotides have other utilities: as intermediates in the biosynthesis of several fine chemicals (e.g., thiamine, S-adenosyl-methionine, folates, or riboflavin), as energy carriers for the cell (e.g., ATP or GTP), and for chemicals themselves, commonly used as flavor enhancers (e.g., IMP or GMP) or for several medicinal applications (see, for example, Kuninaka, A. (1996) Nucleotides and Related Compounds in Biotechnology vol. 6, Rehm et al., eds. VCH: Weinheim, p. 561-612). Also, enzymes involved in purine, pyrimidine, nucleoside, or nucleotide

metabolism are increasingly serving as targets against which chemicals for crop protection, including fungicides, herbicides and insecticides, are developed.

The metabolism of these compounds in bacteria has been characterized (for reviews see, for example, Zalkin, H. and Dixon, J.E. (1992) "*de novo* purine nucleotide biosynthesis", in: Progress in Nucleic Acid Research and Molecular Biology, vol. 42, Academic Press:, p. 259-287; and Michal, G. (1999) "Nucleotides and Nucleosides", Chapter 8 in: Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, Wiley: New York). Purine metabolism has been the subject of intensive research, and is essential to the normal functioning of the cell. Impaired purine metabolism in higher animals can cause severe disease, such as gout. Purine nucleotides are synthesized from ribose-5-phosphate, in a series of steps through the intermediate compound inosine-5'-phosphate (IMP), resulting in the production of guanosine-5'-monophosphate (GMP) or adenosine-5'-monophosphate (AMP), from which the triphosphate forms utilized as nucleotides are readily formed. These compounds are also utilized as energy stores, so their degradation provides energy for many different biochemical processes in the cell. Pyrimidine biosynthesis proceeds by the formation of uridine-5'-monophosphate (UMP) from ribose-5-phosphate. UMP, in turn, is converted to cytidine-5'-triphosphate (CTP). The deoxy- forms of all of these nucleotides are produced in a one step reduction reaction from the diphosphate ribose form of the nucleotide to the diphosphate deoxyribose form of the nucleotide. Upon phosphorylation, these molecules are able to participate in DNA synthesis.

D. Trehalose Metabolism and Uses

Trehalose consists of two glucose molecules, bound in α, α-1,1 linkage. It is commonly used in the food industry as a sweetener, an additive for dried or frozen foods, and in beverages. However, it also has applications in the pharmaceutical, cosmetics and biotechnology industries (see, for example, Nishimoto et al., (1998) U.S. Patent No. 5,759,610; Singer, M.A. and Lindquist, S. (1998) *Trends Biotech.* 16: 460-467; Paiva, C.L.A. and Panek, A.D. (1996) *Biotech. Ann. Rev.* 2: 293-314; and Shiosaka, M. (1997) J. Japan 172: 97-102). Trehalose is produced by enzymes from many microorganisms and is naturally released into the surrounding medium, from which it can be collected using methods known in the art.

II. Mechanisms of Metabolic Regulation

All living cells have complex catabolic and anabolic metabolic capabilities with many interconnected pathways. In order to maintain a balance between the various parts of this extremely complex metabolic network, the cell employs a finely-tuned regulatory

network. By regulating enzyme synthesis and enzyme activity, either independently or simultaneously, the cell is able to control the activity of disparate metabolic pathways to reflect the changing needs of the cell.

The induction or repression of enzyme synthesis may occur at either the level of transcription or translation, or both. Gene expression in prokaryotes is regulated by several mechanisms at the level of transcription (for review see e.g., Lewin, B (1990) Genes IV, Part 3: "Controlling prokaryotic genes by transcription", Oxford University Press: Oxford, p. 213-301, and references therein, and Michal, G. (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley & Sons). All such known regulatory processes are mediated by additional genes, which themselves respond to external influences of various kinds (e.g., temperature, nutrient availability, or light). Exemplary protein factors which have been implicated in this type of regulation include the transcription factors. These are proteins which bind to DNA, thereby either increasing the expression of a gene (positive regulation, as in the case of e.g. the *ara* operon from *E. coli*) or decreasing gene expression (negative regulation, as in the case of the *lac* operon from *E. coli*). These expression-modulating transcription factors can themselves be the subject of regulation. Their activity can, for example, be regulated by the binding of low molecular weight compounds to the DNA-binding protein, thereby stimulating (as in the case of arabinose for the ara operon) or inhibiting (as in the case of the lactose for the lac operon) the binding of these proteins to the appropriate binding site on the DNA (see, for example, Helmann, J.D. and Chamberlin, M.J. (1988) "Structure and function of bacterial sigma factors." *Ann. Rev Biochem.* 57: 839-872; Adhya, S. (1995) "The lac and gal operons today" and Boos, W. et al., "The maltose system.", both in: Regulation of Gene Expression in *Escherichia coli* (Lin, E.C.C. and Lynch, A.S., eds.) Chapman & Hall: New York, p. 181-200 and 201-229; and Moran, C.P. (1993) "RNA polymerase and transcription factors." in: *Bacillus subtilis* and other gram-positive bacteria, Sonenshein, A.L. et al., eds. ASM: Washington, D.C., p. 653-667.)

Aside from the transcriptional level, protein synthesis is also often regulated at the level of translation. There are multiple mechanisms by which such regulation may occur, including alteration of the ability of the ribosome to bind to one or more mRNAs, binding of the ribosome to the mRNA, the maintenance or removal of mRNA secondary structure, the utilization of common or less common codons for a particular gene, the degree of abundance of one or more tRNAs, and special regulation mechanisms, such as attenuation (see Vellanoweth, R.I. (1993) Translation and its regulation in *Bacillus subtilis* and other gram-positive bacteria, Sonenshein, A.L. et al., eds. ASM: Washington, D.C., p. 699-711 and references cited therein).

Transcriptional and translational regulation may be targeted to a single protein (sequential regulation) or simultaneously to several proteins in different metabolic pathways (coordinate regulation). Often, genes whose expression is coordinately regulated are physically located near one another in the genome, in an operon or regulon. Such up- or down-regulation of gene transcription and translation is governed by the cellular and extracellular levels of various factors. such as substrates (precursor and intermediate molecules used in one or more metabolic pathways), catabolites (molecules produced by biochemical pathways concerned with the production of energy from the breakdown of complex organic molecules such as sugars). and end products (the molecules resulting at the end of a metabolic pathway). Typically, the expression of genes encoding enzymes necessary for the activity of a particular pathway is induced by high levels of substrate molecules for that pathway. Similarly. such gene expression tends to be repressed when there exist high intracellular levels of the end product of the pathway (Snyder, L. and Champness, W. (1997) The Molecular Biology of Bacteria ASM: Washington). Gene expression may also be regulated by other external and internal factors, such as environmental conditions (e.g., heat, oxidative stress, or starvation). These global environmental changes cause alterations in the expression of specialized modulating genes, which directly or indirectly (via additional genes or proteins) trigger the expression of genes by means of binding to DNA and thereby inducing or repressing transcription (see. for example, Lin, E.C.C. and Lynch, A.S., eds. (1995) Regulation of Gene Expression in *Escherichia coli*. Chapman & Hall: New York).

Yet another mechanism by which cellular metabolism may be regulated is at the level of the protein. Such regulation is accomplished either by the activities of other proteins, or by binding of low-molecular-weight components which either impede or enable the normal functioning of the protein. Examples of protein regulation by the binding of low-molecular-weight compounds include the binding of GTP or NAD. The binding of a low-molecular-weight chemical is typically reversible, as is the case with the GTP-binding proteins. These proteins exist in two stages (with bound GTP or GDP), one stage being the activated form of the protein. and one stage being inactive.

Regulation of protein activity by the action of other enzymes typically takes the form of covalent modification of the protein (i.e., phosphorylation of amino acid residues such as histidine or aspartate. or methylation). Such covalent modification is typically reversible, as mediated by an enzyme of the opposite activity. An example of this is the opposite activities of kinases and phosphorylases in protein phosphorylation; protein kinases phosphorylate specific residues on a target protein (e.g., serine or threonine). while protein phosphorylases remove phosphate groups from such proteins.

Typically, enzymes which modulate the activity of other proteins are themselves modulated by external stimuli. These stimuli are mediated through proteins which function as sensors. A well known mechanism by which such sensor proteins may mediate these external signals is by dimerization, but others are also known (see, for example, Msadek, T. et al. (1993) "Two-Component Regulatory Systems", in: *Bacillus subtilis* and Other Gram-Positive Bacteria, Sonenshein, A.L. et al., eds., ASM: Washington p. 729-745 and references cited therein).

A thorough understanding of the regulatory networks governing cellular metabolism in microorganisms is critical for the high-yield production of chemicals by fermentation. Control systems for the down-regulation of metabolic pathways could be removed or lessened to improve the synthesis of desired chemicals, and similarly, those for the up-regulation of metabolic pathways for a desired product could be constitutively activated or optimized in activity (As shown in Hirose, Y. and Okada, H. (1979) "Microbial Production of Amino Acids", in: Peppler, H.J. and Perlman, D. (eds.) Microbial Technology 2nd ed. Vol. 1, ch. 7 Academic Press: New York.)

III. Elements and Methods of the Invention

The present invention is based, at least in part, on the discovery of novel molecules, referred to herein as MR nucleic acid and protein molecules, which regulate, by transcriptional, translational, or post-translational means, one or more metabolic pathways in *C. glutamicum*. In one embodiment, the MR molecules transcriptionally, translationally, or posttranslationally regulate a metabolic pathway in *C. glutamicum*. In a preferred embodiment, the activity of the MR molecules of the present invention to regulate one or more *C. glutamicum* metabolic pathways has an impact on the production of a desired fine chemical by this organism. In a particularly preferred embodiment, the MR molecules of the invention are modulated in activity, such that the *C. glutamicum* metabolic pathways which the MR proteins of the invention regulate are modulated in efficiency or output, which either directly or indirectly modulates the yield, production, and/or efficiency of production of a desired fine chemical by *C. glutamicum*.

The language, "MR protein" or "MR polypeptide" includes proteins which transcriptionally, translationally, or posttranslationally regulate a metabolic pathway in *C. glutamicum*. Examples of MR proteins include those encoded by the MR genes set forth in Table 1 and Appendix A. The terms "MR gene" or "MR nucleic acid sequence" include nucleic acid sequences encoding an MR protein, which consist of a coding region and also corresponding untranslated 5' and 3' sequence regions. Examples of MR genes include those set forth in Table 1. The terms "production" or "productivity" are

art-recognized and include the concentration of the fermentation product (for example, the desired fine chemical) formed within a given time and a given fermentation volume (e.g., kg product per hour per liter). The term "efficiency of production" includes the time required for a particular level of production to be achieved (for example, how long it takes for the cell to attain a particular rate of output of a fine chemical). The term "yield" or "product/carbon yield" is art-recognized and includes the efficiency of the conversion of the carbon source into the product (i.e., fine chemical). This is generally written as, for example, kg product per kg carbon source. By increasing the yield or production of the compound, the quantity of recovered molecules, or of useful recovered molecules of that compound in a given amount of culture over a given amount of time is increased. The terms "biosynthesis" or a "biosynthetic pathway" are art-recognized and include the synthesis of a compound, preferably an organic compound, by a cell from intermediate compounds in what may be a multistep and highly regulated process. The terms "degradation" or a "degradation pathway" are art-recognized and include the breakdown of a compound, preferably an organic compound, by a cell to degradation products (generally speaking, smaller or less complex molecules) in what may be a multistep and highly regulated process. The language "metabolism" is art-recognized and includes the totality of the biochemical reactions that take place in an organism. The metabolism of a particular compound, then, (e.g., the metabolism of an amino acid such as glycine) comprises the overall biosynthetic, modification, and degradation pathways in the cell related to this compound. The term, "regulation" is art-recognized and includes the activity of a protein to govern the activity of another protein. The term, "transcriptional regulation" is art-recognized and includes the activity of a protein to impede or activate the conversion of a DNA encoding a target protein to mRNA. The term, "translational regulation" is art-recognized and includes the activity of a protein to impede or activate the conversion of an mRNA encoding a target protein to a protein molecule. The term, "posttranslational regulation" is art-recognized and includes the activity of a protein to impede or improve the activity of a target protein by covalently modifying the target protein (e.g., by methylation, glucosylation, or phosphorylation).

In another embodiment, the MR molecules of the invention are capable of modulating the production of a desired molecule, such as a fine chemical, in a microorganism such as *C. glutamicum*. Using recombinant genetic techniques, one or more of the regulatory proteins of the invention for metabolic pathways may be manipulated such that its function is modulated. For example, a biosynthetic enzyme may be improved in efficiency, or its allosteric control region destroyed such that feedback inhibition of production of the compound is prevented. Similarly, a degradative enzyme may be deleted or modified by substitution, deletion, or addition

such that its degradative activity is lessened for the desired compound without impairing the viability of the cell. In each case, the overall yield or rate of production of one of these desired fine chemicals may be increased.

It is also possible that such alterations in the protein and nucleotide molecules of the invention may improve the production of fine chemicals in an indirect fashion. The regulatory mechanisms of metabolic pathways in the cell are necessarily intertwined, and the activation of one pathway may lead to the repression or activation of another in a concomitant fashion. Therefore, by modulating the activity of one or more of the proteins of the invention, the production or efficiency of activity of another fine chemical biosynthetic or degradative pathway may be impacted. For example, by decreasing the ability of an MR protein to repress the transcription of a gene encoding a particular amino acid biosynthetic protein, one may concomitantly derepress other amino acid biosynthetic pathways, since these pathways are interrelated. Further, by modifying the MR proteins of the invention, one may uncouple the growth and division of cells from their extracellular surroundings to a certain degree; by impairing an MR protein which normally represses biosynthesis of a nucleotide when the extracellular conditions are suboptimal for growth and cell division such that it now lacks this function, one may permit growth to occur even when the extracellular conditions are poor. This is of particular relevance in large-scale fermentative growth, where conditions within the culture are often suboptimal in terms of temperature, nutrient supply or aeration, but would still support growth and cell division if the cellular regulatory systems for these factors were eliminated.

The isolated nucleic acid sequences of the invention are contained within the genome of a *Corynebacterium glutamicum* strain available through the American Type Culture Collection, given designation ATCC 13032. The nucleotide sequence of the isolated *C. glutamicum* MR cDNAs and the predicted amino acid sequences of the *C. glutamicum* MR proteins are shown in Appendices A and B, respectively. Computational analyses were performed which classified and/or identified these nucleotide sequences as sequences which encode metabolic pathway regulatory proteins.

The present invention also pertains to proteins which have an amino acid sequence which is substantially homologous to an amino acid sequence of Appendix B. As used herein, a protein which has an amino acid sequence which is substantially homologous to a selected amino acid sequence is least about 50% homologous to the selected amino acid sequence, e.g., the entire selected amino acid sequence. A protein which has an amino acid sequence which is substantially homologous to a selected amino acid sequence can also be least about 50-60%, preferably at least about 60-70%, and more preferably at least about 70-80%, 80-90%, or 90-95%, and most preferably at

least about 96%, 97%, 98%, 99% or more homologous to the selected amino acid sequence.

The MR protein or a biologically active portion or fragment thereof of the invention can transcriptionally, translationally, or posttranslationally regulate a metabolic pathway in *C glutamicum*, or have one or more of the activities set forth in Table 1.

Various aspects of the invention are described in further detail in the following subsections:

A. Isolated Nucleic Acid Molecules

One aspect of the invention pertains to isolated nucleic acid molecules that encode MR polypeptides or biologically active portions thereof, as well as nucleic acid fragments sufficient for use as hybridization probes or primers for the identification or amplification of MR-encoding nucleic acid (e.g., MR DNA). As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (e.g., cDNA or genomic DNA) and RNA molecules (e.g., mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. This term also encompasses untranslated sequence located at both the 3' and 5' ends of the coding region of the gene: at least about 100 nucleotides of sequence upstream from the 5' end of the coding region and at least about 20 nucleotides of sequence downstream from the 3'end of the coding region of the gene. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA. An "isolated" nucleic acid molecule is one which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated MR nucleic acid molecule can contain less than about 5 kb, 4kb, 3kb, 2kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived (e.g, a *C. glutamicum* cell). Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, e.g., a nucleic acid molecule having a nucleotide sequence of Appendix A, or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. For example, a *C. glutamicum* MR cDNA can be isolated from a *C. glutamicum* library

using all or portion of one of the sequences of Appendix A as a hybridization probe and standard hybridization techniques (e.g., as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd. ed., Cold Spring Harbor Laboratory*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989). Moreover, a nucleic acid molecule encompassing all or a portion of one of the sequences of Appendix A can be isolated by the polymerase chain reaction using oligonucleotide primers designed based upon this sequence (e.g., a nucleic acid molecule encompassing all or a portion of one of the sequences of Appendix A can be isolated by the polymerase chain reaction using oligonucleotide primers designed based upon this same sequence of Appendix A). For example, mRNA can be isolated from normal endothelial cells (e.g., by the guanidinium-thiocyanate extraction procedure of Chirgwin et al. (1979) *Biochemistry* 18: 5294-5299) and cDNA can be prepared using reverse transcriptase (e.g., Moloney MLV reverse transcriptase, available from Gibco/BRL, Bethesda, MD; or AMV reverse transcriptase, available from Seikagaku America, Inc., St. Petersburg, FL). Synthetic oligonucleotide primers for polymerase chain reaction amplification can be designed based upon one of the nucleotide sequences shown in Appendix A. A nucleic acid of the invention can be amplified using cDNA or, alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to an MR nucleotide sequence can be prepared by standard synthetic techniques, e.g., using an automated DNA synthesizer.

In a preferred embodiment, an isolated nucleic acid molecule of the invention comprises one of the nucleotide sequences shown in Appendix A. The sequences of Appendix A correspond to the *Corynebacterium glutamicum* MR cDNAs of the invention. This cDNA comprises sequences encoding MR proteins (i.e., "the coding region", indicated in each sequence in Appendix A), as well as 5' untranslated sequences and 3' untranslated sequences, also indicated in Appendix A. Alternatively, the nucleic acid molecule can comprise only the coding region of any of the sequences in Appendix A.

For the purposes of this application, it will be understood that each of the sequences set forth in Appendix A has an identifying RXA number having the designation "RXA" followed by 5 digits (i.e., RXA00004). Each of these sequences comprises up to three parts: a 5' upstream region, a coding region, and a downstream region. Each of these three regions is identified by the same RXA designation to eliminate confusion. The recitation "one of the sequences in Appendix A", then, refers to any of the sequences in Appendix A, which may be distinguished by their differing

RXA designations. The coding region of each of these sequences is translated into a corresponding amino acid sequence, which is set forth in Appendix B. The sequences of Appendix B are identified by the same RXA designations as Appendix A, such that they can be readily correlated. For example, the amino acid sequence in Appendix B designated RXA00004 is a translation of the coding region of the nucleotide sequence of nucleic acid molecule RXA00004 in Appendix A.

In one embodiment, the nucleic acid molecules of the present invention are not intended to include those compiled in Table 2. In the case of the dapD gene, a sequence for this gene was published in Wehrmann, A., et al. (1998) *J Bacteriol.* 180(12): 3159-3165. However, the sequence obtained by the inventors of the present application is significantly longer than the published version. It is believed that the published version relied on an incorrect start codon, and thus represents only a fragment of the actual coding region.

In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of one of the nucleotide sequences shown in Appendix A, or a portion thereof. A nucleic acid molecule which is complementary to one of the nucleotide sequences shown in Appendix A is one which is sufficiently complementary to one of the nucleotide sequences shown in Appendix A such that it can hybridize to one of the nucleotide sequences shown in Appendix A, thereby forming a stable duplex.

In still another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleotide sequence which is at least about 50-60%, preferably at least about 60-70%, more preferably at least about 70-80%, 80-90%, or 90-95%, and even more preferably at least about 95%, 96%, 97%, 98%, 99% or more homologous to a nucleotide sequence shown in Appendix A, or a portion thereof. In an additional preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleotide sequence which hybridizes, e.g., hybridizes under stringent conditions, to one of the nucleotide sequences shown in Appendix A, or a portion thereof.

Moreover, the nucleic acid molecule of the invention can comprise only a portion of the coding region of one of the sequences in Appendix A, for example a fragment which can be used as a probe or primer or a fragment encoding a biologically active portion of an MR protein. The nucleotide sequences determined from the cloning of the MR genes from *C. glutamicum* allows for the generation of probes and primers designed for use in identifying and/or cloning MR homologues in other cell types and organisms, as well as MR homologues from other *Corynebacteria* or related species. The probe/primer typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes

under stringent conditions to at least about 12, preferably about 25, more preferably about 40, 50 or 75 consecutive nucleotides of a sense strand of one of the sequences set forth in Appendix A, an anti-sense sequence of one of the sequences set forth in Appendix A, or naturally occurring mutants thereof. Primers based on a nucleotide sequence of Appendix A can be used in PCR reactions to clone MR homologues. Probes based on the MR nucleotide sequences can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In preferred embodiments, the probe further comprises a label group attached thereto, e.g. the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells which misexpress an MR protein, such as by measuring a level of an MR-encoding nucleic acid in a sample of cells, e.g., detecting MR mRNA levels or determining whether a genomic MR gene has been mutated or deleted.

In one embodiment, the nucleic acid molecule of the invention encodes a protein or portion thereof which includes an amino acid sequence which is sufficiently homologous to an amino acid sequence of Appendix B such that the protein or portion thereof maintains the ability to transcriptionally, translationally, or posttranslationally regulate a metabolic pathway in *C. glutamicum*. As used herein, the language "sufficiently homologous" refers to proteins or portions thereof which have amino acid sequences which include a minimum number of identical or equivalent (e.g., an amino acid residue which has a similar side chain as an amino acid residue in one of the sequences of Appendix B) amino acid residues to an amino acid sequence of Appendix B such that the protein or portion thereof is able to transcriptionally, translationally, or posttranslationally regulate a metabolic pathway in *C. glutamicum*. Protein members of such metabolic pathways, as described herein, may function to regulate the biosynthesis or degradation of one or more fine chemicals. Examples of such activities are also described herein. Thus, "the function of an MR protein" contributes to the overall regulation of one or more fine chemical metabolic pathway. Examples of MR protein activities are set forth in Table 1.

In another embodiment, the protein is at least about 50-60%, preferably at least about 60-70%, and more preferably at least about 70-80%, 80-90%, 90-95%, and most preferably at least about 96%, 97%, 98%, 99% or more homologous to an entire amino acid sequence of Appendix B.

Portions of proteins encoded by the MR nucleic acid molecules of the invention are preferably biologically active portions of one of the MR proteins. As used herein, the term "biologically active portion of an MR protein" is intended to include a portion, e.g., a domain/motif, of an MR protein that transcriptionally, translationally, or

posttranslationally regulates a metabolic pathway in *C. glutamicum*, or has an activity as set forth in Table 1. To determine whether an MR protein or a biologically active portion thereof can transcriptionally, translationally, or posttranslationally regulate a metabolic pathway in *C. glutamicum*, an assay of enzymatic activity may be performed. Such assay methods are well known to those skilled in the art, as detailed in Example 8 of the Exemplification.

Additional nucleic acid fragments encoding biologically active portions of an MR protein can be prepared by isolating a portion of one of the sequences in Appendix B, expressing the encoded portion of the MR protein or peptide (e.g., by recombinant expression *in vitro*) and assessing the activity of the encoded portion of the MR protein or peptide.

The invention further encompasses nucleic acid molecules that differ from one of the nucleotide sequences shown in Appendix A (and portions thereof) due to degeneracy of the genetic code and thus encode the same MR protein as that encoded by the nucleotide sequences shown in Appendix A. In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in Appendix B. In a still further embodiment, the nucleic acid molecule of the invention encodes a full length *C. glutamicum* protein which is substantially homologous to an amino acid sequence of Appendix B (encoded by an open reading frame shown in Appendix A).

In addition to the *C. glutamicum* MR nucleotide sequences shown in Appendix A, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of MR proteins may exist within a population (e.g., the *C. glutamicum* population). Such genetic polymorphism in the MR gene may exist among individuals within a population due to natural variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame encoding an MR protein, preferably a *C. glutamicum* MR protein. Such natural variations can typically result in 1-5% variance in the nucleotide sequence of the MR gene. Any and all such nucleotide variations and resulting amino acid polymorphisms in MR that are the result of natural variation and that do not alter the functional activity of MR proteins are intended to be within the scope of the invention.

Nucleic acid molecules corresponding to natural variants and non-*C. glutamicum* homologues of the *C. glutamicum* MR cDNA of the invention can be isolated based on their homology to the *C. glutamicum* MR nucleic acid disclosed herein using the *C. glutamicum* cDNA, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions. Accordingly, in

another embodiment, an isolated nucleic acid molecule of the invention is at least 15 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising a nucleotide sequence of Appendix A. In other embodiments, the nucleic acid is at least 30, 50, 100, 250 or more nucleotides in length. As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% homologous to each other typically remain hybridized to each other. Preferably, the conditions are such that sequences at least about 65%, more preferably at least about 70%, and even more preferably at least about 75% or more homologous to each other typically remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. A preferred, non-limiting example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65°C. Preferably, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to a sequence of Appendix A corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (e.g., encodes a natural protein). In one embodiment, the nucleic acid encodes a natural *C. glutamicum* MR protein.

In addition to naturally-occurring variants of the MR sequence that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into a nucleotide sequence of Appendix A, thereby leading to changes in the amino acid sequence of the encoded MR protein, without altering the functional ability of the MR protein. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made in a sequence of Appendix A. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of one of the MR proteins (Appendix B) without altering the activity of said MR protein, whereas an "essential" amino acid residue is required for MR protein activity. Other amino acid residues, however, (e.g., those that are not conserved or only semi-conserved in the domain having MR activity) may not be essential for activity and thus are likely to be amenable to alteration without altering MR activity.

Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding MR proteins that contain changes in amino acid residues that are not essential for MR activity. Such MR proteins differ in amino acid sequence from a sequence contained in Appendix B yet retain at least one of the MR activities described herein. In

one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 50% homologous to an amino acid sequence of Appendix B and is capable of transcriptionally, translationally, or posttranslationally regulating a metabolic pathway in *C. glutamicum*, or has one or more activities set forth in Table 1. Preferably, the protein encoded by the nucleic acid molecule is at least about 50-60% homologous to one of the sequences in Appendix B, more preferably at least about 60-70% homologous to one of the sequences in Appendix B, even more preferably at least about 70-80%, 80-90%, 90-95% homologous to one of the sequences in Appendix B, and most preferably at least about 96%, 97%, 98%, or 99% homologous to one of the sequences in Appendix B.

To determine the percent homology of two amino acid sequences (e.g., one of the sequences of Appendix B and a mutant form thereof) or of two nucleic acids, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of one protein or nucleic acid for optimal alignment with the other protein or nucleic acid). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in one sequence (e.g., one of the sequences of Appendix B) is occupied by the same amino acid residue or nucleotide as the corresponding position in the other sequence (e.g., a mutant form of the sequence selected from Appendix B), then the molecules are homologous at that position (i.e., as used herein amino acid or nucleic acid "homology" is equivalent to amino acid or nucleic acid "identity"). The percent homology between the two sequences is a function of the number of identical positions shared by the sequences (i.e., % homology = # of identical positions/total # of positions x 100).

An isolated nucleic acid molecule encoding an MR protein homologous to a protein sequence of Appendix B can be created by introducing one or more nucleotide substitutions, additions or deletions into a nucleotide sequence of Appendix A such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced into one of the sequences of Appendix A by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine,

tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in an MR protein is preferably replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of an MR coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for an MR activity described herein to identify mutants that retain MR activity. Following mutagenesis of one of the sequences of Appendix A, the encoded protein can be expressed recombinantly and the activity of the protein can be determined using, for example, assays described herein (see Example 8 of the Exemplification).

In addition to the nucleic acid molecules encoding MR proteins described above, another aspect of the invention pertains to isolated nucleic acid molecules which are antisense thereto. An "antisense" nucleic acid comprises a nucleotide sequence which is complementary to a "sense" nucleic acid encoding a protein, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire MR coding strand, or to only a portion thereof. In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding an MR protein. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues (e.g., the entire coding region of SEQ ID RXA00004 comprises nucleotides 1 to 471). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding MR. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (i.e., also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding MR disclosed herein (e.g., the sequences set forth in Appendix A), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of MR mRNA, but more preferably is an oligonucleotide which is antisense to only a portion of the coding or noncoding region of MR mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of MR mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed

using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a cell or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding an MR protein to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. The antisense molecule can be modified such that it specifically binds to a receptor or an antigen expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecule to a peptide or an antibody which binds to a cell surface receptor or antigen. The antisense nucleic acid molecule can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gaultier et al. (1987) *Nucleic Acids Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue et al. (1987) *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue et al. (1987) *FEBS Lett.* 215:327-330).

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave MR mRNA transcripts to thereby inhibit translation of MR mRNA. A ribozyme having specificity for an MR-encoding nucleic acid can be designed based upon the nucleotide sequence of an MR cDNA disclosed herein (i.e., RXA00004 in Appendix A). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in an MR-encoding mRNA. See, e.g., Cech et al. U.S. Patent No. 4,987,071 and Cech et al. U.S. Patent No. 5,116,742. Alternatively, MR mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel, D. and Szostak, J.W. (1993) *Science* 261:1411-1418.

Alternatively, MR gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of an MR nucleotide sequence (e.g., an MR promoter and/or enhancers) to form triple helical structures that prevent transcription of an MR gene in target cells. See generally, Helene, C. (1991) *Anticancer Drug Des.* 6(6):569-84; Helene, C. et al. (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher, L.J. (1992) *Bioassays* 14(12):807-15.

B. Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding an MR protein (or a portion thereof). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of

autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (e.g., in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (e.g., polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cell and those which direct expression of the nucleotide sequence only in certain host cells. It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (e.g., MR proteins, mutant forms of MR proteins, fusion proteins, etc.).

The recombinant expression vectors of the invention can be designed for expression of MR proteins in prokaryotic or eukaryotic cells. For example, MR genes can be expressed in bacterial cells such as *C. glutamicum*, insect cells (using baculovirus expression vectors), yeast and other fungal cells (see Romanos, M.A. et al. (1992)

"Foreign gene expression in yeast: a review", *Yeast* 8: 423-488; van den Hondel, C.A.M.J.J. et al. (1991) "Heterologous gene expression in filamentous fungi" in: More Gene Manipulations in Fungi. J.W. Bennet & L.L. Lasure, eds., p. 396-428: Academic Press: San Diego; and van den Hondel, C.A.M.J.J. & Punt, P.J. (1991) "Gene transfer systems and vector development for filamentous fungi, in: Applied Molecular Genetics of Fungi, Peberdy, J.F. et al., eds., p. 1-28, Cambridge University Press: Cambridge), algae and multicellular plant cells (see Schmidt, R. and Willmitzer, L. (1988) High efficiency *Agrobacterium tumefactiens* –mediated transformation of *Arabidopsis thaliana* leaf and cotyledon explants" *Plant Cell Rep.*: 583-586), or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase.

Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith, D.B. and Johnson, K.S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein. In one embodiment, the coding sequence of the MR protein is cloned into a pGEX expression vector to create a vector encoding a fusion protein comprising, from the N-terminus to the C-terminus, GST-thrombin cleavage site-X protein. The fusion protein can be purified by affinity chromatography using glutathione-agarose resin. Recombinant MR protein unfused to GST can be recovered by cleavage of the fusion protein with thrombin.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann et al., (1988) *Gene* 69:301-315) and pET 11d (Studier et al., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego,

California (1990) 60-89). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident λ prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter.

One strategy to maximize recombinant protein expression is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in the bacterium chosen for expression, such as *C glutamicum* (Wada et al. (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the MR protein expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerivisae* include pYepSec1 (Baldari, et al., (1987) *Embo J.* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz et al., (1987) *Gene* 54:113-123), and pYES2 (Invitrogen Corporation, San Diego, CA). Vectors and methods for the construction of vectors appropriate for use in other fungi, such as the filamentous fungi, include those detailed in: van den Hondel, C.A.M.J.J. & Punt, P.J. (1991) "Gene transfer systems and vector development for filamentous fungi, in: Applied Molecular Genetics of Fungi, J.F. Peberdy, et al., eds., p. 1-28, Cambridge University Press: Cambridge.

Alternatively, the MR proteins of the invention can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (e.g., Sf 9 cells) include the pAc series (Smith et al. (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

In another embodiment, the MR proteins of the invention may be expressed in unicellular plant cells (such as algae) or in plant cells from higher plants (e.g., the spermatophytes, such as crop plants). Examples of plant expression vectors include those detailed in: Becker, D., Kemper, E., Schell, J. and Masterson, R. (1992) "New plant binary vectors with selectable markers located proximal to the left border", *Plant Mol. Biol.* 20: 1195-1197; and Bevan, M.W. (1984) "Binary *Agrobacterium* vectors for plant transformation". *Nucl. Acid. Res.* 12: 8711-8721.

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, B. (1987) *Nature* 329:840) and pMT2PC (Kaufman et al. (1987) *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd. ed., Cold Spring Harbor Laboratory*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (e.g., tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert et al. (1987) *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji et al. (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (e.g., the neurofilament promoter; Byrne and Ruddle (1989) *PNAS* 86:5473-5477), pancreas-specific promoters (Edlund et al. (1985) *Science* 230:912-916), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the α-fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev.* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to MR mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid

or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub, H. et al., Antisense RNA as a molecular tool for genetic analysis, *Reviews - Trends in Genetics*, Vol. 1(1) 1986.

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, an MR protein can be expressed in bacterial cells such as *C. glutamicum*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art. Microorganisms related to *Corynebacterium glutamicum* which may be conveniently used as host cells for the nucleic acid and protein molecules of the invention are set forth in Table 3.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (e.g . DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, et al. (*Molecular Cloning: A Laboratory Manual. 2nd. ed., Cold Spring Harbor Laboratory*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (e.g., resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding an MR protein or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic

acid can be identified by, for example, drug selection (e.g., cells that have incorporated the selectable marker gene will survive, while the other cells die).

To create a homologous recombinant microorganism, a vector is prepared which contains at least a portion of an MR gene into which a deletion, addition or substitution has been introduced to thereby alter, e.g., functionally disrupt, the MR gene. Preferably, this MR gene is a *Corynebacterium glutamicum* MR gene, but it can be a homologue from a related bacterium or even from a mammalian, yeast, or insect source. In a preferred embodiment, the vector is designed such that, upon homologous recombination, the endogenous MR gene is functionally disrupted (i.e., no longer encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous MR gene is mutated or otherwise altered but still encodes functional protein (e.g., the upstream regulatory region can be altered to thereby alter the expression of the endogenous MR protein). In the homologous recombination vector, the altered portion of the MR gene is flanked at its 5' and 3' ends by additional nucleic acid of the MR gene to allow for homologous recombination to occur between the exogenous MR gene carried by the vector and an endogenous MR gene in a microorganism. The additional flanking MR nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector (see e.g., Thomas, K.R., and Capecchi, M.R. (1987) Cell 51: 503 for a description of homologous recombination vectors). The vector is introduced into a microorganism (e.g., by electroporation) and cells in which the introduced MR gene has homologously recombined with the endogenous MR gene are selected, using art-known techniques.

In another embodiment, recombinant microorganisms can be produced which contain selected systems which allow for regulated expression of the introduced gene. For example, inclusion of an MR gene on a vector placing it under control of the lac operon permits expression of the MR gene only in the presence of IPTG. Such regulatory systems are well known in the art.

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (i.e., express) an MR protein. Accordingly, the invention further provides methods for producing MR proteins using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding an MR protein has been introduced, or into which genome has been introduced a gene encoding a wild-type or altered MR protein) in a suitable medium until MR protein is produced. In another

embodiment, the method further comprises isolating MR proteins from the medium or the host cell.

C. Isolated MR Proteins

Another aspect of the invention pertains to isolated MR proteins, and biologically active portions thereof. An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material when produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of MR protein in which the protein is separated from cellular components of the cells in which it is naturally or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of MR protein having less than about 30% (by dry weight) of non-MR protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-MR protein, still more preferably less than about 10% of non-MR protein, and most preferably less than about 5% non-MR protein. When the MR protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, i.e., culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation. The language "substantially free of chemical precursors or other chemicals" includes preparations of MR protein in which the protein is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of MR protein having less than about 30% (by dry weight) of chemical precursors or non-MR chemicals, more preferably less than about 20% chemical precursors or non-MR chemicals, still more preferably less than about 10% chemical precursors or non-MR chemicals, and most preferably less than about 5% chemical precursors or non-MR chemicals. In preferred embodiments, isolated proteins or biologically active portions thereof lack contaminating proteins from the same organism from which the MR protein is derived. Typically, such proteins are produced by recombinant expression of, for example, a *C. glutamicum* MR protein in a microorganism such as *C. glutamicum*.

An isolated MR protein or a portion thereof of the invention can transcriptionally, translationally, or posttranslationally regulate a metabolic pathway in *C. glutamicum*, or has one or more of the activities set forth in Table 1. In preferred embodiments, the protein or portion thereof comprises an amino acid sequence which is sufficiently homologous to an amino acid sequence of Appendix B such that the protein

or portion thereof maintains the ability to transcriptionally, translationally, or posttranslationally regulate a metabolic pathway in *C. glutamicum*. The portion of the protein is preferably a biologically active portion as described herein. In another preferred embodiment, an MR protein of the invention has an amino acid sequence shown in Appendix B. In yet another preferred embodiment, the MR protein has an amino acid sequence which is encoded by a nucleotide sequence which hybridizes, e.g., hybridizes under stringent conditions, to a nucleotide sequence of Appendix A. In still another preferred embodiment, the MR protein has an amino acid sequence which is encoded by a nucleotide sequence that is at least about 50-60%, preferably at least about 60-70%, more preferably at least about 70-80%, 80-90%, 90-95%, and even more preferably at least about 96%, 97%, 98%, 99% or more homologous to one of the amino acid sequences of Appendix B. The preferred MR proteins of the present invention also preferably possess at least one of the MR activities described herein. For example, a preferred MR protein of the present invention includes an amino acid sequence encoded by a nucleotide sequence which hybridizes, e.g., hybridizes under stringent conditions, to a nucleotide sequence of Appendix A, and which can transcriptionally, translationally, or posttranslationally regulate a metabolic pathway in *C. glutamicum*, or which has one or more of the activities set forth in Table 1.

In other embodiments, the MR protein is substantially homologous to an amino acid sequence of Appendix B and retains the functional activity of the protein of one of the sequences of Appendix B yet differs in amino acid sequence due to natural variation or mutagenesis, as described in detail in subsection I above. Accordingly, in another embodiment, the MR protein is a protein which comprises an amino acid sequence which is at least about 50-60%, preferably at least about 60-70%, and more preferably at least about 70-80, 80-90, 90-95%, and most preferably at least about 96%, 97%, 98%, 99% or more homologous to an entire amino acid sequence of Appendix B and which has at least one of the MR activities described herein. In another embodiment, the invention pertains to a full length *C. glutamicum* protein which is substantially homologous to an entire amino acid sequence of Appendix B.

Biologically active portions of an MR protein include peptides comprising amino acid sequences derived from the amino acid sequence of an MR protein, e.g., the an amino acid sequence shown in Appendix B or the amino acid sequence of a protein homologous to an MR protein, which include fewer amino acids than a full length MR protein or the full length protein which is homologous to an MR protein, and exhibit at least one activity of an MR protein. Typically, biologically active portions (peptides, e.g., peptides which are, for example, 5, 10, 15, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 or more amino acids in length) comprise a domain or motif with at least one activity of an

MR protein. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the activities described herein. Preferably, the biologically active portions of an MR protein include one or more selected domains/motifs or portions thereof having biological activity.

MR proteins are preferably produced by recombinant DNA techniques. For example, a nucleic acid molecule encoding the protein is cloned into an expression vector (as described above), the expression vector is introduced into a host cell (as described above) and the MR protein is expressed in the host cell. The MR protein can then be isolated from the cells by an appropriate purification scheme using standard protein purification techniques. Alternative to recombinant expression, an MR protein, polypeptide, or peptide can be synthesized chemically using standard peptide synthesis techniques. Moreover, native MR protein can be isolated from cells (e.g., endothelial cells), for example using an anti-MR antibody, which can be produced by standard techniques utilizing an MR protein or fragment thereof of this invention.

The invention also provides MR chimeric or fusion proteins. As used herein, an MR "chimeric protein" or "fusion protein" comprises an MR polypeptide operatively linked to a non-MR polypeptide. An "MR polypeptide" refers to a polypeptide having an amino acid sequence corresponding to an MR protein, whereas a "non-MR polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein which is not substantially homologous to the MR protein, e.g., a protein which is different from the MR protein and which is derived from the same or a different organism. Within the fusion protein, the term "operatively linked" is intended to indicate that the MR polypeptide and the non-MR polypeptide are fused in-frame to each other. The non-MR polypeptide can be fused to the N-terminus or C-terminus of the MR polypeptide. For example, in one embodiment the fusion protein is a GST-MR fusion protein in which the MR sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of recombinant MR proteins. In another embodiment, the fusion protein is an MR protein containing a heterologous signal sequence at its N-terminus. In certain host cells (e.g., mammalian host cells), expression and/or secretion of an MR protein can be increased through use of a heterologous signal sequence.

Preferably, an MR chimeric or fusion protein of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini.

filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel et al. John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). An MR-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the MR protein.

Homologues of the MR protein can be generated by mutagenesis, e.g., discrete point mutation or truncation of the MR protein. As used herein, the term "homologue" refers to a variant form of the MR protein which acts as an agonist or antagonist of the activity of the MR protein. An agonist of the MR protein can retain substantially the same, or a subset, of the biological activities of the MR protein. An antagonist of the MR protein can inhibit one or more of the activities of the naturally occurring form of the MR protein, by, for example, competitively binding to a downstream or upstream member of the MR regulatory cascade which includes the MR protein. Thus, the *C. glutamicum* MR protein and homologues thereof of the present invention may modulate the activity of one or more metabolic pathways which MR proteins regulate in this microorganism.

In an alternative embodiment, homologues of the MR protein can be identified by screening combinatorial libraries of mutants, e.g., truncation mutants, of the MR protein for MR protein agonist or antagonist activity. In one embodiment, a variegated library of MR variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of MR variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential MR sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g., for phage display) containing the set of MR sequences therein. There are a variety of methods which can be used to produce libraries of potential MR homologues from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential MR sequences. Methods for synthesizing degenerate

oligonucleotides are known in the art (see, e.g., Narang, S.A. (1983) *Tetrahedron* 39:3; Itakura et al. (1984) *Annu. Rev. Biochem.* 53:323; Itakura et al. (1984) *Science* 198:1056; Ike et al. (1983) *Nucleic Acid Res.* 11:477.

In addition, libraries of fragments of the MR protein coding can be used to generate a variegated population of MR fragments for screening and subsequent selection of homologues of an MR protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of an MR coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal, C-terminal and internal fragments of various sizes of the MR protein.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of MR homologues. The most widely used techniques, which are amenable to high through-put analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recrusive ensemble mutagenesis (REM), a new technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify MR homologues (Arkin and Yourvan (1992) *PNAS* 89:7811-7815; Delgrave et al. (1993) *Protein Engineering* 6(3):327-331).

In another embodiment, cell based assays can be exploited to analyze a variegated MR library, using methods well known in the art.

D. Uses and Methods of the Invention

The nucleic acid molecules, proteins, protein homologues, fusion proteins, primers, vectors, and host cells described herein can be used in one or more of the following methods: identification of *C. glutamicum* and related organisms; mapping of genomes of organisms related to *C. glutamicum*; identification and localization of *C. glutamicum* sequences of interest; evolutionary studies; determination of MR protein regions required for function; modulation of an MR protein activity; modulation of the

activity of one or more metabolic pathways; and modulation of cellular production of a desired compound, such as a fine chemical.

The MR nucleic acid molecules of the invention have a variety of uses. First, they may be used to identify an organism as being *Corynebacterium glutamicum* or a close relative thereof. Also, they may be used to identify the presence of *C. glutamicum* or a relative thereof in a mixed population of microorganisms. The invention provides the nucleic acid sequences of a number of *C. glutamicum* genes; by probing the extracted genomic DNA of a culture of a unique or mixed population of microorganisms under stringent conditions with a probe spanning a region of a *C. glutamicum* gene which is unique to this organism, one can ascertain whether this organism is present. Although *Corynebacterium glutamicum* itself is nonpathogenic, it is related to pathogenic species, such as *Corynebacterium diphtheriae*. Detection of such organisms is of significant clinical relevance.

Further, the nucleic acid and protein molecules of the invention may serve as markers for specific regions of the genome. This has utility not only in the mapping of the genome, but also for functional studies of *C. glutamicum* proteins. For example, to identify the region of the genome to which a particular *C. glutamicum* DNA-binding protein binds, the *C. glutamicum* genome could be digested, and the fragments incubated with the DNA-binding protein. Those which bind the protein may be additionally probed with the nucleic acid molecules of the invention, preferably with readily detectable labels; binding of such a nucleic acid molecule to the genome fragment enables the localization of the fragment to the genome map of *C. glutamicum*, and, when performed multiple times with different enzymes, facilitates a rapid determination of the nucleic acid sequence to which the protein binds. Further, the nucleic acid molecules of the invention may be sufficiently homologous to the sequences of related species such that these nucleic acid molecules may serve as markers for the construction of a genomic map in related bacteria, such as *Brevibacterium lactofermentum*.

The MR nucleic acid molecules of the invention are also useful for evolutionary and protein structural studies. The metabolic processes in which the molecules of the invention participate are utilized by a wide variety of prokaryotic and eukaryotic cells; by comparing the sequences of the nucleic acid molecules of the present invention to those encoding similar enzymes from other organisms, the evolutionary relatedness of the organisms can be assessed. Similarly, such a comparison permits an assessment of which regions of the sequence are conserved and which are not, which may aid in determining those regions of the protein which are essential for the functioning of the enzyme. This type of determination is of value for protein engineering studies and may

give an indication of what the protein can tolerate in terms of mutagenesis without losing function.

Manipulation of the MR nucleic acid molecules of the invention may result in the production of MR proteins having functional differences from the wild-type MR proteins. These proteins may be improved in efficiency or activity, may be present in greater numbers in the cell than is usual, or may be decreased in efficiency or activity.

Such changes in activity may directly modulate the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*. For example, by optimizing the activity of an MR protein which activates the transcription or translation of a gene encoding a biosynthetic protein for a desired fine chemical, or by impairing or abrogating the activity of an MR protein which represses the transcription or translation of such a gene, one may also increase the activity or rate of activity of that biosynthetic pathway due to the presence of increased levels of what may have been a limiting enzyme. Similarly, by altering the activity of an MR protein such that it constitutively posttranslationally inactivates a protein involved in a degradation pathway for a desired fine chemical, or by altering the activity of an MR protein such that it constitutively represses the transcription or translation of such a gene, one may increase the yield and/or rate of production of the fine chemical from the cell, due to decreased degradation of the compound.

Further, by modulating the activity of one or more MR proteins, one may indirectly stimulate the production or improve the rate of production of one or more fine chemicals from the cell due to the interrelatedness of disparate metabolic pathways. For example, by increasing the yield, production, and/or efficiency of production by activating the expression of one or more lysine biosynthetic enzymes, one may concomitantly increase the expression of other compounds, such as other amino acids, which the cell would naturally require in greater quantities when lysine is required in greater quantities. Also, regulation of metabolism throughout the cell may be altered such that the cell is better able to grow or replicate under the environmental conditions of fermentative culture (where nutrient and oxygen supplies may be poor and possibly toxic waste products in the environment may be at high levels). For example, by mutagenizing an MR protein which represses the synthesis of molecules necessary for cell membrane production in response to high levels of waste products in the extracellular medium (in order to block cell growth and division in suboptimal growth conditions) such that it no longer is able to repress such synthesis, one may increase the growth and multiplication of the cell in cultures even when the growth conditions are suboptimal. Such enhanced growth or viability should also increase the yields and/or

rate of production of a desired fine chemical from fermentative culture, due to the relatively greater number of cells producing this compound in the culture.

The aforementioned mutagenesis strategies for MR proteins to result in increased yields of a fine chemical from *C. glutamicum* are not meant to be limiting; variations on these strategies will be readily apparent to one skilled in the art. Using such strategies, and incorporating the mechanisms disclosed herein, the nucleic acid and protein molecules of the invention may be utilized to generate *C. glutamicum* or related strains of bacteria expressing mutated MR nucleic acid and protein molecules such that the yield and/or efficiency of production of a desired compound is improved. This desired compound may be any natural product of *C. glutamicum*, which includes the final products of biosynthesis pathways and intermediates of naturally-occurring metabolic pathways, as well as molecules which do not naturally occur in the metabolism of *C glutamicum*, but which are produced by a *C glutamicum* strain of the invention.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patent applications, patents, and published patent applications cited throughout this application are hereby incorporated by reference.

Exemplification

Example 1: Preparation of total genomic DNA of *Corynebacterium glutamicum* ATCC 13032

A culture of *Corynebacterium glutamicum* (ATCC 13032) was grown overnight at 30°C with vigorous shaking in BHI medium (Difco). The cells were harvested by centrifugation, the supernatant was discarded and the cells were resuspended in 5 ml buffer-I (5% of the original volume of the culture — all indicated volumes have been calculated for 100 ml of culture volume). Composition of buffer-I: 140.34 g/l sucrose, 2.46 g/l $MgSO_4 \times 7H_2O$, 10 ml/l KH_2PO_4 solution (100 g/l, adjusted to pH 6.7 with KOH), 50 ml/l M12 concentrate (10 g/l $(NH_4)_2SO_4$, 1 g/l NaCl, 2 g/l $MgSO_4 \times 7H_2O$, 0.2 g/l $CaCl_2$, 0.5 g/l yeast extract (Difco), 10 ml/l trace-elements-mix (200 mg/l $FeSO_4 \times H_2O$, 10 mg/l $ZnSO_4 \times 7 H_2O$, 3 mg/l $MnCl_2 \times 4 H_2O$, 30 mg/l H_3BO_3, 20 mg/l $CoCl_2 \times 6 H_2O$, 1 mg/l $NiCl_2 \times 6 H_2O$, 3 mg/l $Na_2MoO_4 \times 2 H_2O$, 500 mg/l complexing agent (EDTA or critic acid), 100 ml/l vitamins-mix (0.2 mg/l biotin, 0.2 mg/l folic acid, 20 mg/l p-amino benzoic acid, 20 mg/l riboflavin, 40 mg/l ca-panthothenate, 140 mg/l nicotinic acid, 40 mg/l pyridoxole hydrochloride, 200 mg/l myo-inositol). Lysozyme was added to the suspension to a final concentration of 2.5 mg/ml. After an approximately 4 h incubation at 37°C, the cell wall was degraded and the resulting protoplasts are harvested by centrifugation. The pellet was washed once with 5 ml buffer-I and once with 5 ml TE-buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8). The pellet was resuspended in 4 ml TE-buffer and 0.5 ml SDS solution (10%) and 0.5 ml NaCl solution (5 M) are added. After adding of proteinase K to a final concentration of 200 µg/ml, the suspension is incubated for ca.18 h at 37°C. The DNA was purified by extraction with phenol, phenol-chloroform-isoamylalcohol and chloroform-isoamylalcohol using standard procedures. Then, the DNA was precipitated by adding 1/50 volume of 3 M sodium acetate and 2 volumes of ethanol, followed by a 30 min incubation at -20°C and a 30 min centrifugation at 12,000 rpm in a high speed centrifuge using a SS34 rotor (Sorvall). The DNA was dissolved in 1 ml TE-buffer containing 20

µg/ml RNaseA and dialysed at 4°C against 1000 ml TE-buffer for at least 3 hours. During this time, the buffer was exchanged 3 times. To aliquots of 0.4 ml of the dialysed DNA solution, 0.4 ml of 2 M LiCl and 0.8 ml of ethanol are added. After a 30 min incubation at -20°C, the DNA was collected by centrifugation (13,000 rpm, Biofuge Fresco, Heraeus, Hanau, Germany). The DNA pellet was dissolved in TE-buffer. DNA prepared by this procedure could be used for all purposes, including southern blotting or construction of genomic libraries.

Example 2: Construction of genomic libraries in *Escherichia coli* of *Corynebacterium glutamicum* ATCC13032.

Starting from DNA prepared as described in Example 1, cosmid and plasmid libraries were constructed according to known and well established methods (*see* e.g., Sambrook, J. *et al* (1989) "Molecular Cloning : A Laboratory Manual", Cold Spring Harbor Laboratory Press, or Ausubel, F.M. *et al.* (1994) "Current Protocols in Molecular Biology", John Wiley & Sons.)

Any plasmid or cosmid could be used. Of particular use were the plasmids pBR322 (Sutcliffe, J.G. (1979) Proc. Natl. Acad. Sci. USA, 75:3737-3741); pACYC177 (Change & Cohen (1978) J. Bacteriol 134:1141-1156), plasmids of the pBS series (pBSSK+, pBSSK- and others; Stratagene, LaJolla, USA), or cosmids as SuperCos1 (Stratagene, LaJolla, USA) or Lorist6 (Gibson, T.J., Rosenthal A. and Waterson, R.H. (1987) Gene 53:283-286.

Example 3: DNA Sequencing and Computational Functional Analysis

Genomic libraries as described in Example 2 were used for DNA sequencing according to standard methods, in particular by the chain termination method using ABI377 sequencing machines (see e.g., Fleischman, R.D. et al. (1995) "Whole-genome Random Sequencing and Assembly of Haemophilus Influenzae Rd., Science, 269:496-512). Sequencing primers with the following nucleotide sequences were used: 5`-GGAAACAGTATGACCATG-3` or 5`-GTAAAACGACGGCCAGT-3`.

Example 4: *In vivo* Mutagenesis

In vivo mutagenesis of *Corynebacterium glutamicum* can be performed by passage of plasmid (or other vector) DNA through *E. coli* or other microorganisms (e.g. *Bacillus* spp. or yeasts such as *Saccharomyces cerevisiae*) which are impaired in their capabilities to maintain

the integrity of their genetic information. Typical mutator strains have mutations in the genes for the DNA repair system (e.g., mutHLS, mutD, mutT, etc.: for reference, see Rupp, W.D. (1996) DNA repair mechanisms, in: *Escherichia coli* and *Salmonella*, p. 2277-2294, ASM: Washington.) Such strains are well known to those skilled in the art. The use of such strains is illustrated, for example, in Greener, A. and Callahan, M. (1994) Strategies 7: 32-34.

Example 5: DNA Transfer Between *Escherichia coli* and *Corynebacterium glutamicum*

Several *Corynebacterium* and *Brevibacterium* species contain endogenous plasmids (as e.g., pHM1519 or pBL1) which replicate autonomously (for review see, e.g., Martin, J.F. *et al.* (1987) Biotechnology, 5:137-146). Shuttle vectors for *Escherichia coli* and *Corynebacterium glutamicum* can be readily constructed by using standard vectors for *E coli* (Sambrook, J. *et al.* (1989). "Molecular Cloning: A Laboratory Manual", Cold Spring Harbor Laboratory Press or Ausubel, F.M. *et al.* (1994) "Current Protocols in Molecular Biology". John Wiley & Sons) to which a origin or replication for and a suitable marker from *Corynebacterium glutamicum* is added. Such origins of replication are preferably taken from endogenous plasmids isolated from *Corynebacterium* and *Brevibacterium* species. Of particular use as transformation markers for these species are genes for kanamycin resistance (such as those derived from the Tn5 or Tn903 transposons) or chloramphenicol (Winnacker, E.L. (1987) "From Genes to Clones — Introduction to Gene Technology, VCH, Weinheim). There are numerous examples in the literature of the construction of a wide variety of shuttle vectors which replicate in both *E. coli and C glutamicum*, and which can be used for several purposes, including gene over-expression (for reference, see e.g., Yoshihama, M. et al. (1985) J. Bacteriol. 162:591-597, Martin J.F. et al. (1987) Biotechnology, 5:137-146 and Eikmanns, B.J. et al. (1991) Gene, 102:93-98).

Using standard methods, it is possible to clone a gene of interest into one of the shuttle vectors described above and to introduce such a hybrid vectors into strains of *Corynebacterium glutamicum*. Transformation of *C. glutamicum* can be achieved by protoplast transformation (Kastsumata, R. et al. (1984) J. Bacteriol. 159306-311), electroporation (Liebl, E. et al. (1989) FEMS Microbiol. Letters, 53:399-303) and in cases where special vectors are used, also by conjugation (as described e.g. in Schäfer, A et al. (1990) J. Bacteriol. 172:1663-1666). It is also possible to transfer the shuttle vectors for

C. glutamicum to *E. coli* by preparing plasmid DNA from *C. glutamicum* (using standard methods well-known in the art) and transforming it into *E. coli*. This transformation step can be performed using standard methods, but it is advantageous to use an Mcr-deficient *E. coli* strain, such as NM522 (Gough & Murray (1983) J. Mol. Biol. 166:1-19).

Example 6: Assessment of the Expression of the Mutant Protein

Observations of the activity of a mutated protein in a transformed host cell rely on the fact that the mutant protein is expressed in a similar fashion and in a similar quantity to that of the wild-type protein. A useful method to ascertain the level of transcription of the mutant gene (an indicator of the amount of mRNA available for translation to the gene product) is to perform a Northern blot (for reference see, for example, Ausubel et al. (1988) Current Protocols in Molecular Biology, Wiley: New York), in which a primer designed to bind to the gene of interest is labeled with a detectable tag (usually radioactive or chemiluminescent), such that when the total RNA of a culture of the organism is extracted, run on gel, transferred to a stable matrix and incubated with this probe, the binding and quantity of binding of the probe indicates the presence and also the quantity of mRNA for this gene. This information is evidence of the degree of transcription of the mutant gene. Total cellular RNA can be prepared from *Corynebacterium glutamicum* by several methods, all well-known in the art, such as that described in Bormann, E.R. et al. (1992) *Mol. Microbiol.* 6: 317-326.

To assess the presence or relative quantity of protein translated from this mRNA, standard techniques, such as a Western blot, may be employed (see, for example, Ausubel et al. (1988) Current Protocols in Molecular Biology, Wiley: New York). In this process, total cellular proteins are extracted, separated by gel electrophoresis, transferred to a matrix such as nitrocellulose, and incubated with a probe, such as an antibody, which specifically binds to the desired protein. This probe is generally tagged with a chemiluminescent or colorimetric label which may be readily detected. The presence and quantity of label observed indicates the presence and quantity of the desired mutant protein present in the cell.

Example 7: Growth of Genetically Modified *Corynebacterium glutamicum* — Media and Culture Conditions

Genetically modified *Corynebacteria* are cultured in synthetic or natural growth media. A number of different growth media for Corynebacteria are both well-known and readily available (Lieb *et al.* (1989) Appl. Microbiol. Biotechnol., 32:205-210; von der Osten *et al* (1998) Biotechnology Letters. 11:11-16; Patent DE 4,120.867; Liebl (1992) "The Genus *Corynebacterium*, in: The Procaryotes, Volume II, Balows, A. *et al.*, eds. Springer-Verlag). These media consist of one or more carbon sources, nitrogen sources, inorganic salts, vitamins and trace elements. Preferred carbon sources are sugars, such as mono-, di-, or polysaccharides. For example, glucose, fructose, mannose, galactose, ribose, sorbose, ribulose, lactose, maltose, sucrose, raffinose, starch or cellulose serve as very good carbon sources. It is also possible to supply sugar to the media via complex compounds such as molasses or other by-products from sugar refinement. It can also be advantageous to supply mixtures of different carbon sources. Other possible carbon sources are alcohols and organic acids, such as methanol, ethanol, acetic acid or lactic acid. Nitrogen sources are usually organic or inorganic nitrogen compounds, or materials which contain these compounds. Exemplary nitrogen sources include ammonia gas or ammonia salts, such as NH_4Cl or $(NH_4)_2SO_4$, NH_4OH, nitrates, urea, amino acids or complex nitrogen sources like corn steep liquor, soy bean flour, soy bean protein, yeast extract, meat extract and others.

Inorganic salt compounds which may be included in the media include the chloride-, phosphorous- or sulfate- salts of calcium, magnesium, sodium, cobalt, molybdenum, potassium, manganese, zinc, copper and iron. Chelating compounds can be added to the medium to keep the metal ions in solution. Particularly useful chelating compounds include dihydroxyphenols, like catechol or protocatechuate, or organic acids, such as citric acid. It is typical for the media to also contain other growth factors, such as vitamins or growth promoters, examples of which include biotin, riboflavin, thiamin, folic acid, nicotinic acid, pantothenate and pyridoxin. Growth factors and salts frequently originate from complex media components such as yeast extract, molasses, corn steep liquor and others. The exact composition of the media compounds depends strongly on the immediate experiment and is individually decided for each specific case. Information about media optimization is available in the textbook "Applied Microbiol. Physiology, A Practical Approach (*eds.* P.M. Rhodes, P.F. Stanbury, IRL Press (1997) pp. 53-73, ISBN 0

19 963577 3). It is also possible to select growth media from commercial suppliers, like standard 1 (Merck) or BHI (grain heart infusion, DIFC) or others.

All medium components are sterilized, either by heat (20 minutes at 1.5 bar and 121°C) or by sterile filtration. The components can either be sterilized together or, if necessary, separately. All media components can be present at the beginning of growth or they can optionally be added continuously or batchwise.

Culture conditions are defined separately for each experiment. The temperature should be in a range between 15°C and 45°C. The temperature can be kept constant or can be altered during the experiment. The pH of the medium should be in the range of 5 to 8.5, preferably around 7.0, and can be maintained by the addition of buffers to the media. An exemplary buffer for this purpose is a potassium phosphate buffer. Synthetic buffers such as MOPS, HEPES, ACES and others can alternatively or simultaneously be used. It is also possible to maintain a constant culture pH through the addition of NaOH or NH_4OH during growth. If complex medium components such as yeast extract are utilized, the necessity for additional buffers may be reduced, due to the fact that many complex compounds have high buffer capacities. If a fermentor is utilized for culturing the micro-organisms, the pH can also be controlled using gaseous ammonia.

The incubation time is usually in a range from several hours to several days. This time is selected in order to permit the maximal amount of product to accumulate in the broth. The disclosed growth experiments can be carried out in a variety of vessels, such as microtiter plates, glass tubes, glass flasks or glass or metal fermentors of different sizes. For screening a large number of clones, the microorganisms should be cultured in microtiter plates, glass tubes or shake flasks, either with or without baffles. Preferably 100 ml shake flasks are used, filled with 10% (by volume) of the required growth medium. The flasks should be shaken on a rotary shaker (amplitude 25 mm) using a speed-range of 100 – 300 rpm. Evaporation losses can be diminished by the maintenance of a humid atmosphere; alternatively, a mathematical correction for evaporation losses should be performed.

If genetically modified clones are tested, an unmodified control clone or a control clone containing the basic plasmid without any insert should also be tested. The medium is inoculated to an OD_{600} of 0.5 – 1.5 using cells grown on agar plates, such as CM plates (10 g/l glucose, 2.5 g/l NaCl, 2 g/l urea, 10 g/l polypeptone, 5 g/l yeast extract, 5 g/l meat extract, 22 g/l NaCl, 2 g/l urea, 10 g/l polypeptone, 5 g/l yeast extract, 5 g/l meat extract.

22 g/l agar, pH 6.8 with 2M NaOH) that had been incubated at 30°C. Inoculation of the media is accomplished by either introduction of a saline suspension of *C. glutamicum* cells from CM plates or addition of a liquid preculture of this bacterium.

Example 8 – *In vitro* Analysis of the Function of Mutant Proteins

The determination of activities and kinetic parameters of enzymes is well established in the art. Experiments to determine the activity of any given altered enzyme must be tailored to the specific activity of the wild-type enzyme, which is well within the ability of one skilled in the art. Overviews about enzymes in general, as well as specific details concerning structure, kinetics, principles, methods, applications and examples for the determination of many enzyme activities may be found, for example, in the following references: Dixon, M., and Webb, E.C., (1979) Enzymes. Longmans: London; Fersht, (1985) Enzyme Structure and Mechanism. Freeman: New York; Walsh, (1979) Enzymatic Reaction Mechanisms. Freeman: San Francisco; Price, N.C., Stevens, L. (1982) Fundamentals of Enzymology. Oxford Univ. Press: Oxford; Boyer, P.D., ed. (1983) The Enzymes, 3rd ed. Academic Press: New York; Bisswanger, H., (1994) Enzymkinetik, 2nd ed. VCH: Weinheim (ISBN 3527300325); Bergmeyer, H.U., Bergmeyer, J., Graßl, M., eds. (1983-1986) Methods of Enzymatic Analysis, 3rd ed., vol. I-XII, Verlag Chemie: Weinheim; and Ullmann's Encyclopedia of Industrial Chemistry (1987) vol. A9, "Enzymes". VCH: Weinheim, p. 352-363.

The activity of proteins which bind to DNA can be measured by several well-established methods, such as DNA band-shift assays (also called gel retardation assays). The effect of such proteins on the expression of other molecules can be measured using reporter gene assays (such as that described in Kolmar, H. et al. (1995) EMBO J. 14: 3895-3904 and references cited therein). Reporter gene test systems are well known and established for applications in both pro- and eukaryotic cells, using enzymes such as beta-galactosidase, green fluorescent protein, and several others.

The determination of activity of membrane-transport proteins can be performed according to techniques such as those described in Gennis, R.B. (1989) "Pores, Channels and Transporters", in Biomembranes, Molecular Structure and Function. Springer: Heidelberg, p. 85-137; 199-234; and 270-322.

Example 9: Analysis of Impact of Mutant Protein on the Production of the Desired Product

The effect of the genetic modification in *C glutamicum* on production of a desired compound (such as an amino acid) can be assessed by growing the modified microorganism under suitable conditions (such as those described above) and analyzing

the medium and/or the cellular component for increased production of the desired product (i.e., an amino acid). Such analysis techniques are well known to one skilled in the art, and include spectroscopy, thin layer chromatography, staining methods of various kinds, enzymatic and microbiological methods, and analytical chromatography such as high performance liquid chromatography (see, for example, Ullman, Encyclopedia of Industrial Chemistry, vol. A2, p. 89-90 and p. 443-613. VCH: Weinheim (1985); Fallon, A. et al., (1987) "Applications of HPLC in Biochemistry" in: Laboratory Techniques in Biochemistry and Molecular Biology, vol. 17; Rehm et al. (1993) Biotechnology, vol. 3, Chapter III: "Product recovery and purification", page 469-714. VCH: Weinheim: Belter, P.A. et al. (1988) Bioseparations: downstream processing for biotechnology, John Wiley and Sons; Kennedy, J.F. and Cabral, J.M.S. (1992) Recovery processes for biological materials, John Wiley and Sons; Shaeiwitz, J.A. and Henry, J.D. (1988) Biochemical separations, in: Ullmann's Encyclopedia of Industrial Chemistry, vol. B3, Chapter 11, page 1-27, VCH: Weinheim; and Dechow, F.J. (1989) Separation and purification techniques in biotechnology, Noyes Publications.)

In addition to the measurement of the final product of fermentation, it is also possible to analyze other components of the metabolic pathways utilized for the production of the desired compound, such as intermediates and side-products, to determine the overall yield, production, and/or efficiency of production of the compound. Analysis methods include measurements of nutrient levels in the medium (e.g., sugars, hydrocarbons, nitrogen sources, phosphate, and other ions), measurements of biomass composition and growth, analysis of the production of common metabolites of biosynthetic pathways, and measurement of gasses produced during fermentation. Standard methods for these measurements are outlined in Applied Microbial Physiology, A Practical Approach, P.M. Rhodes and P.F. Stanbury, eds., IRL Press, p. 103-129; 131-163; and 165-192 (ISBN: 0199635773) and references cited therein.

Example 10: Purification of the Desired Product from *C. glutamicum* Culture

Recovery of the desired product from the *C. glutamicum* cells or supernatant of the above-described culture can be performed by various methods well known in the art. If the desired product is not secreted from the cells, the cells can be harvested from the culture by low-speed centrifugation, the cells can be lysed by standard techniques, such as mechanical force or sonication. The cellular debris is removed by centrifugation, and the supernatant fraction containing the soluble proteins is retained for further purification of the desired compound. If the product is secreted from the *C. glutamicum*

cells, then the cells are removed from the culture by low-speed centrifugation, and the supernate fraction is retained for further purification.

The supernatant fraction from either purification method is subjected to chromatography with a suitable resin, in which the desired molecule is either retained on a chromatography resin while many of the impurities in the sample are not, or where the impurities are retained by the resin while the sample is not. Such chromatography steps may be repeated as necessary, using the same or different chromatography resins. One skilled in the art would be well-versed in the selection of appropriate chromatography resins and in their most efficacious application for a particular molecule to be purified. The purified product may be concentrated by filtration or ultrafiltration, and stored at a temperature at which the stability of the product is maximized.

There are a wide array of purification methods known to the art and the preceding method of purification is not meant to be limiting. Such purification techniques are described, for example, in Bailey, J.E. & Ollis, D.F. Biochemical Engineering Fundamentals, McGraw-Hill: New York (1986).

The identity and purity of the isolated compounds may be assessed by techniques standard in the art. These include high-performance liquid chromatography (HPLC), spectroscopic methods, staining methods, thin layer chromatography, NIRS, enzymatic assay, or microbiologically. Such analysis methods are reviewed in: Patek et al. (1994) *Appl. Environ. Microbiol.* 60: 133-140; Malakhova et al. (1996) *Biotekhnologiya* 11: 27-32; and Schmidt et al. (1998) *Bioprocess Engineer.* 19: 67-70. Ulmann's Encyclopedia of Industrial Chemistry, (1996) vol. A27, VCH: Weinheim, p. 89-90, p. 521-540, p. 540-547, p. 559-566, 575-581 and p. 581-587; Michal, G. (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley and Sons; Fallon, A. et al. (1987) Applications of HPLC in Biochemistry in: Laboratory Techniques in Biochemistry and Molecular Biology, vol. 17.

Equivalents

Those skilled in the art will recognize, or will be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

TABLE 1: GENES INCLUDED IN THE APPLICATION

Genes for signal transduction pathways, regulation of proteins and transcription

ID #	Contig	NT Start	NT Stop	Name of Gene	Function of Gene
RXA00291	GR00047	2	1075	BS-cssS,EC-b0619	SENSOR KINASE CITA (EC 2 7 3 -)
RXA00129	GR00020	6205	4709		SENSOR PROTEIN CPXA (EC 2 7 3 -)
RXA00651	GR00169	5450	4119	BS-yvqE	SENSOR PROTEIN DEGS (EC 2.7 3 -)
RXA00006	GR00001	6905	6471		SENSOR PROTEIN FIXL (EC 2 7 3 -)
RXA01860	GR00529	2368	1484	EC-b2503,BS-ykoW	SENSOR PROTEIN FIXL (EC 2 7 3 -)
RXA01861	GR00529	4332	2368	EC-yegE	SENSOR PROTEIN FIXL (EC 2 7 3 -)
RXA02669	GR00753	8893	10008	EC-baeS	SENSOR PROTEIN RESE (EC 2 7 3 -)
RXA01211	GR00349	741	1535		SENSOR PROTEIN UHPB (EC 2 7 3 -)
RXA01248	GR00362	165	593		SENSORY TRANSDUCTION PROTEIN REGX3
RXA02668	GR00753	8171	8893		SENSORY TRANSDUCTION PROTEIN REGX3
RXA01139	GR00318	2138	3091	BS-czcD,EC-b0752	membrane-bound sensor for cations
RXA02632	GR00748	4863	4168	BS-ykoG	putative two-component response regulator [Mycobacterium tuberculosis]
RXA02631	GR00748	4096	2732		putative two-component sensor [Mycobacterium tuberculosis]
RXA00609	GR00161	226	891	EC-b0571	TWO COMPONENT RESPONSE REGULATOR
RXA02391	GR00695	1894	1508		(U59457) Pseudomonas aeruginosa ankyrin (ankB) gene, complete cds [Pseudomonas aeruginosa]
RXA00284	GR00045	1318	2382		ANKYRIN HOMOLOG PRECURSOR
RXA01152	GR00324	1641	907		PROTEIN-L-ISOASPARTATE O-METHYLTRANSFERASE (EC 2 1 1 77)
RXA01827	GR00516	6308	4902		PROTEIN KINASE PKNA
RXA00813	GR00219	1345	2475		SECRETORY PROTEIN KINASE
RXA01826	GR00516	4902	2965	BS-yloP	PUTATIVE SERINE/THREONINE-PROTEIN KINASE PKNB (EC 2.7.1 -)
RXA02699	GR00757	1357	3504		PUTATIVE SERINE/THREONINE-PROTEIN KINASE PKNB (EC 2 7 1 -)
RXA00319	GR00056	505	80		LOW MOLECULAR WEIGHT PHOSPHOTYROSINE PROTEIN PHOSPHATASE (EC 3.1.3 48)
RXA01272	GR00367	25049	24447		PROBABLE LOW MOLECULAR WEIGHT PROTEIN-TYROSINE-PHOSPHATASE EPSP (EC 3 1.3 48)
RXA01830	GR00516	10410	9058	BS-yloO	PUTATIVE PHOSPHOPROTEIN PHOSPHATASE
RXA02747	GR00764	277	2352		[PROTEIN-PII] URIDYLYLTRANSFERASE (EC 2 7 7 59)
RXA00786	GR00208	3	881	EC-lytB,BS-yqfP	LYTB PROTEIN, amidase enhancer
RXA02210	GR00648	1922	2485		Hypothetical Transcriptional Regulator
RXA00221	GR00032	20855	21073		Hypothetical Transcriptional Regulator
RXA00551	GR00144	352	5		Hypothetical Transcriptional Regulator
RXA01763	GR00500	1987	1523		Hypothetical Transcriptional Regulator
RXA02667	GR00753	7863	7270	BS-ydeB	Hypothetical Transcriptional Regulator
RXA00348	GR00065	1507	1052		Hypothetical Transcriptional Regulator
RXA01500	GR00424	7551	7108		Hypothetical Transcriptional Regulator
RXA01125	GR00312	1800	1588	BS-yozG	Hypothetical Transcriptional Regulator

RXA00822	GR00221	3073	2393	EC-crp,BS-fnr	putative transcriptional regulator
RXA00849	GR00231	378	698		possible transcriptional regulator
RXA02698	GR00757	1143	775		PUTATIVE TRANSCRIPTIONAL REGULATOR
RXA00350	GR00066	1144	1470		Hypothetical Transcription Initiation Factor
RXA02830	GR00817	3	497		Helix-turn-helix domain-containing transcription regulators
RXA00947	GR00259	4164	3829	BS-yczG	Helix-turn-helix domain-containing transcriptional regulators
RXA02732	GR00762	5493	5933		Putative transcription factors
RXA01835	GR00517	4370	3666		(AL021287) probable transcriptional regulator [Mycobacterium tuberculosis]
RXA00292	GR00047	1078	1731	EC-cirR,BS-ydbG	transcriptional regulator CirR
RXA00182	GR00028	4247	7348	BS-ywqA	POSSIBLE GLOBAL TRANSCRIPTION ACTIVATOR SNF2L
RXA02760	GR00767	1154	201	BS-nusG,EC-nusG	TRANSCRIPTION ANTITERMINATION PROTEIN NUSG
RXA02306	GR00663	3214	2924		TRANSCRIPTIONAL REGULATORY PROTEIN CITB
RXA00130	GR00020	6985	6308	BS-ycbL,EC-phoB	TRANSCRIPTIONAL REGULATORY PROTEIN CPXR
RXA00885	GR00242	11301	12326	BS-hrcA	HEAT-INDUCIBLE TRANSCRIPTION REPRESSOR HRCA
RXA02880	GR10018	417	4		TRANSCRIPTIONAL REPRESSOR CYTR
RXA01418	GR00415	776	531		TRANSCRIPTIONAL REPRESSOR SMTB
RXA01404	GR00410	7	999		TRANSCRIPTION REGULATORY PROTEIN PEPR1
RXA01759	GR00498	4075	4836		TRANSCRIPTIONAL REGULATORY PROTEIN GLTC
RXA02289	GR00661	1272	1631		SIGMA B TRANSCRIBED GENE
RXA00033	GR00003	12891	12076	BS-yoaB	SIGMA B TRANSCRIBED GENE
RXA00363	GR00073	1929	1246		NTA OPERON TRANSCRIPTIONAL REGULATOR
RXA00516	GR00131	592	1311		NTA OPERON TRANSCRIPTIONAL REGULATOR
RXA01537	GR00427	4829	4179		NTA OPERON TRANSCRIPTIONAL REGULATOR
RXA02494	GR00720	4169	4864	BS-phoP	KDP OPERON TRANSCRIPTIONAL REGULATORY PROTEIN KDPE
RXA02763	GR00768	1603	2586	EC-fhuR	MALTOSE OPERON TRANSCRIPTIONAL REPRESSOR
RXA00029	GR00003	8910	8374		PUTATIVE AGA OPERON TRANSCRIPTIONAL REPRESSOR
RXA00655	GR00169	9049	8411		putative regulatory protein
RXA00645	GR00168	5831	8161		PUTATIVE REGULATORY PROTEIN
RXA00593	GR00158	2858	2511		REGULATORY PROTEIN
RXA02724	GR00760	870	4		REGULATORY PROTEIN
RXA00494	GR00123	768	472		Hypothetical Regulatory Protein
RXA01368	GR00397	2334	2206		Hypothetical Regulatory Protein
RXA00464	GR00117	75	332		REGULATORY PROTEIN SIR2 HOMOLOG .
RXA01655	GR00480	1458	100	BS-ydeL	PROBABLE RHIZOPINE CATABOLISM REGULATORY PROTEIN MOCR
RXA00126	GR00020	2269	1607		PROBABLE SIGMA(54) MODULATION PROTEIN
RXA02450	GR00710	2533	3087		POTENTIAL ACRAB OPERON REPRESSOR
RXA01898	GR00544	1178	1870		OPERON REGULATOR
RXA00004	GR00001	4293	3823		NITRILASE REGULATOR
RXA01001	GR00284	516	833		hex regulon repressor hexR
RXA01375	GR00400	2560	1106		FRNA
RXA00784	GR00206	6452	6024		FRNE
RXA00603	GR00159	4982	5434	EC-lrp,BS-azlB	LEUCINE-RESPONSIVE REGULATORY PROTEIN
RXA01533	GR00426	3877	4617	BS-yktC,EC-suhB	EXTRAGENIC SUPPRESSOR PROTEIN SUHB
RXA02831	GR00818	411	4		EXTRAGENIC SUPPRESSOR PROTEIN SUHB
RXA01110	GR00306	16399	16971		TETRACYCLINE REPRESSOR PROTEIN CLASS C
RXA00253	GR00038	1064	1801		TETRACYCLINE REPRESSOR PROTEIN CLASS E
RXA02493	GR00720	2931	4169	BS-yyoG,EC-phoR	regulatory gene for the phosphate regulon
RXA01118	GR00309	1787	2551	EC-iclR	regulator of the glyoxylate bypass
RXA01840	GR00521	2	655		ALIPHATIC AMIDASE EXPRESSION-REGULATING PROTEIN

RXA00307	GR00052	467	6		DIPHTHERIA TOXIN REPRESSOR
RXA00400	GR00087	1163	2041	EC-b1595,BS-alsR	ALS OPERON REGULATORY PROTEIN
RXA02787	GR00777	865	2241	EC-ycaJ,BS-yrvN	ACTIVATOR 1 41 KD SUBUNIT
RXA00287	GR00046	1618	1145	BS-dat	ADAPTIVE RESPONSE REGULATORY PROTEIN
RXA01386	GR00406	6102	5476		NITRILASE REGULATOR
RXA01687	GR00470	3289	2219		N-ACETYLGLUCOSAMINE REPRESSOR
RXA01935	GR00555	8902	7739		N-ACETYLGLUCOSAMINE REPRESSOR
RXA02270	GR00655	5005	4385		member of the regulatory protein family SIR2
RXA01241	GR00359	739	1218	BS-lexA,EC-lexA	LEXA REPRESSOR (EC 3.4.21.88)
RXA00818	GR00220	624	4		INHIBITION OF MORPHOLOGICAL DIFFERENTIATION
RXA02127	GR00637	2715	2062		6 ACTVA REGION GENES OF THE ACTINORHODIN BIOSYNTHETIC GENE CLUSTER
RXA00583	GR00156	10203	9466	BS-yluA	Uncharacterized ACR (translation?)
RXA00592	GR00158	2121	1663		Uncharacterized ACR (translation initiation regulator?)
RXA00630	GR00166	2	160		(U67196) DNA-binding response regulator [Thermotoga maritima]
RXA00638	GR00167	2862	3245		DNA-binding response regulator
RXA00894	GR00244	1926	799	EC-b0581	GTPASE-ACTIVATING PROTEIN 1
RXA01450	GR00419	1237	1800	BS-yyaF,EC-ychF	GTP-BINDING PROTEIN
RXA01451	GR00419	1760	2326		GTP-BINDING PROTEIN
RXA02376	GR00689	3064	1562	BS-obg,EC-yhbZ	GTP-BINDING PROTEIN
RXA01065	GR00298	2	583	BS-bex,EC-era	GTP-BINDING PROTEIN ERA
RXA02232	GR00653	5286	6812	EC-hflX,BS-ynbA	GTP-BINDING PROTEIN HFLX
RXA00839	GR00228	372	4		GTP-BINDING PROTEIN LEPA
RXA00845	GR00229	907	5	EC-yihK,BS-ylaG	GTP-BINDING PROTEIN LEPA
RXA00848	GR00230	2125	1955		GTP-BINDING PROTEIN LEPA
RXA02365	GR00686	1568	1029		GTP-BINDING PROTEIN LEPA
RXA02392	GR00696	1264	5	BS-lepA,EC-lepA	GTP-BINDING PROTEIN LEPA
RXA01573	GR00438	5744	3663	BS-yhcR	2',3'-cyclic-nucleotide 2'-posphodiesterase

TABLE 2: GENES IDENTIFIED FROM GENBANK

GenBank™ Accession No.	Gene Name	Gene Function	Reference
A09073	ppg	Phosphoenol pyruvate carboxylase	Bachmann, B. et al. "DNA fragment coding for phosphoenolpyruvat corboxylase, recombinant DNA carrying said fragment, strains carrying the recombinant DNA and method for producing L-aminino acids using said strains," Patent: EP 0358940-A 3 03/21/90
A45579, A45581, A45583, A45585 A45587		Threonine dehydratase	Moeckel, B. et al. "Production of L-isoleucine by means of recombinant micro-organisms with deregulated threonine dehydratase," Patent. WO 9519442-A 5 07/20/95
AB003132	murC, ftsQ; ftsZ		Kobayashi, M. et al. "Cloning, sequencing, and characterization of the ftsZ gene from coryneform bacteria," *Biochem. Biophys. Res. Commun.*, 236(2):383-388 (1997)
AB015023	murC; ftsQ		Wachi, M. et al. "A murC gene from Coryneform bacteria," *Appl. Microbiol. Biotechnol*, 51(2):223-228 (1999)
AB018530	dtsR		Kimura, E. et al. "Molecular cloning of a novel gene, dtsR, which rescues the detergent sensitivity of a mutant derived from *Brevibacterium lactofermentum*," *Biosci. Biotechnol. Biochem*, 60(10).1565-1570 (1996)
AB018531	dtsR1; dtsR2		
AB020624	murI	D-glutamate racemase	
AB023377	tkt	transketolase	
AB024708	gltB, gltD	Glutamine 2-oxoglutarate aminotransferase large and small subunits	
AB025424	acn	aconitase	
AB027714	rep	Replication protein	
AB027715	rep; aad	Replication protein; aminoglycoside adenyltransferase	
AF005242	argC	N-acetylglutamate-5-semialdehyde dehydrogenase	
AF005635	glnA	Glutamine synthetase	
AF030405	hisF	cyclase	
AF030520	argG	Argininosuccinate synthetase	
AF031518	argF	Ornithine carbamolytransferase	
AF036932	aroD	3-dehydroquinate dehydratase	

GenBank™ Accession No.	Gene Name	Gene Function	Reference
AF038548	pyc	Pyruvate carboxylase	
AF038651	dciAE; apt; rel	Dipeptide-binding protein; adenine phosphoribosyltransferase; GTP pyrophosphokinase	Wehmeier, L. et al. "The role of the Corynebacterium glutamicum rel gene in (p)ppGpp metabolism," *Microbiology*, 144.1853-1862 (1998)
AF041436	argR	Arginine repressor	
AF045998	impA	Inositol monophosphate phosphatase	
AF048764	argH	Argininosuccinate lyase	
AF049897	argC; argJ; argB; argD; argF; argR; argG; argH	N-acetylglutamylphosphate reductase, ornithine acetyltransferase; N-acetylglutamate kinase, acetylornithine transminase; ornithine carbamoyltransferase; arginine repressor; argininosuccinate synthase; argininosuccinate lyase	
AF050109	inhA	Enoyl-acyl carrier protein reductase	
AF050166	hisG	ATP phosphoribosyltransferase	
AF051846	hisA	Phosphoribosylforminino-5-amino-1-phosphoribosyl-4-imidazolecarboxamide isomerase	
AF052652	metA	Homoserine O-acetyltransferase	Park, S. et al. "Isolation and analysis of metA, a methionine biosynthetic gene encoding homoserine acetyltransferase in Corynebacterium glutamicum," *Mol Cells.*, 8(3):286-294 (1998)
AF053071	aroB	Dehydroquinate synthetase	
AF060558	hisH	Glutamine amidotransferase	
AF086704	hisE	Phosphoribosyl-ATP-pyrophosphohydrolase	
AF114233	aroA	5-enolpyruvylshikimate 3-phosphate synthase	
AF116184	panD	L-aspartate-alpha-decarboxylase precursor	Dusch, N. et al. "Expression of the Corynebacterium glutamicum panD gene encoding L-aspartate-alpha-decarboxylase leads to pantothenate overproduction in Escherichia coli," *Appl. Environ Microbiol.*, 65(4)1530-1539 (1999)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
AF124518	aroD; aroE	3-dehydroquinase; shikimate dehydrogenase	
AF124600	aroC; aroK; aroB; pepQ	Chorismate synthase; shikimate kinase; 3-dehydroquinate synthase; putative cytoplasmic peptidase	
AF145897	inhA		
AF145898	inhA		
AJ001436	ectP	Transport of ectoine, glycine betaine, proline	Peter, H. et al. "Corynebacterium glutamicum is equipped with four secondary carriers for compatible solutes. Identification, sequencing, and characterization of the proline/ectoine uptake system, ProP, and the ectoine/proline/glycine betaine carrier, EctP," *J. Bacteriol.*, 180(22):6005-6012 (1998)
AJ004934	dapD	Tetrahydrodipicolinate succinylase (incomplete')	Wehrmann, A. et al. "Different modes of diaminopimelate synthesis and their role in cell wall integrity: A study with Corynebacterium glutamicum," *J Bacteriol.*, 180(12):3159-3165 (1998)
AJ007732	ppc; secG; amt; ocd; soxA	Phosphoenolpyruvate-carboxylase; ?; high affinity ammonium uptake protein; putative ornithine-cyclodecarboxylase; sarcosine oxidase	
AJ010319	ftsY, glnB, glnD; srp; amtP	Involved in cell division; PII protein; uridylyltransferase (uridylyl-removing enzanye); signal recognition particle; low affinity ammonium uptake protein	Jakoby, M. et al. "Nitrogen regulation in Corynebacterium glutamicum; Isolation of genes involved in biochemical characterization of corresponding proteins," *FEMS Microbiol*, 173(2):303-310 (1999)
AJ132968	cat	Chloramphenicol acetcyl transferase	
AJ224946	mqo	L-malate: quinone oxidoreductase	Molenaar, D. et al. "Biochemical and genetic characterization of the membrane-associated malate dehydrogenase (acceptor) from Corynebacterium glutamicum," *Eur. J Biochem*, 254(2):395-403 (1998)
AJ238250	ndh	NADH dehydrogenase	
AJ238703	porA	Porin	Lichtinger, T. et al "Biochemical and biophysical characterization of the cell wall porin of Corynebacterium glutamicum: The channel is formed by a low molecular mass polypeptide," *Biochemistry*, 37(43):15024-15032 (1998)
D17429		Transposable element IS31831	Vertes, A.A. et al. "Isolation and characterization of IS31831, a transposable element from Corynebacterium glutamicum," *Mol. Microbiol.*, 11(4):739-746 (1994)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
D84102	odhA	2-oxoglutarate dehydrogenase	Usuda, Y. et al. "Molecular cloning of the Corynebacterium glutamicum (Brevibacterium lactofermentum AJ12036) odhA gene encoding a novel type of 2-oxoglutarate dehydrogenase," *Microbiology*, 142.3347-3354 (1996)
E01358	hdh, hk	Homoserine dehydrogenase; homoserine kinase	Katsumata, R. et al. "Production of L-threonine and L-isoleucine," Patent: JP 1987232392-A 1 10/12/87
E01359		Upstream of the start codon of homoserine kinase gene	Katsumata, R. et al. "Production of L-threonine and L-isoleucine," Patent: JP 1987232392-A 2 10/12/87
E01375		Tryptophan operon	
E01376	trpL; trpE	Leader peptide; anthranilate synthase	Matsui, K. et al. "Tryptophan operon, peptide and protein coded thereby, utilization of tryptophan operon gene expression and production of tryptophan," Patent: JP 1987244382-A 1 10/24/87
E01377		Promoter and operator regions of tryptophan operon	Matsui, K. et al. "Tryptophan operon, peptide and protein coded thereby, utilization of tryptophan operon gene expression and production of tryptophan," Patent: JP 1987244382-A 1 10/24/87
E03937		Biotin-synthase	Hatakeyama, K. et al "DNA fragment containing gene capable of coding biotin synthetase and its utilization," Patent: JP 1992278088-A 1 10/02/92
E04040		Diamino pelargonic acid aminotransferase	Kohama, K. et al. "Gene coding diaminopelargonic acid aminotransferase and desthiobiotin synthetase and its utilization," Patent: JP 1992330284-A 1 11/18/92
E04041		Desthiobiotinsynthetase	Kohama, K. et al. "Gene coding diaminopelargonic acid aminotransferase and desthiobiotin synthetase and its utilization," Patent: JP 1992330284-A 1 11/18/92
E04307		Flavum aspartase	Kurusu, Y. et al "Gene DNA coding aspartase and utilization thereof," Patent: JP 1993030977-A 1 02/09/93
E04376		Isocitric acid lyase	Katsumata, R. et al. "Gene manifestation controlling DNA," Patent: JP 1993056782-A 3 03/09/93
E04377		Isocitric acid lyase N-terminal fragment	Katsumata, R et al. "Gene manifestation controlling DNA," Patent. JP 1993056782-A 3 03/09/93
E04484		Prephenate dehydratase	Sotouchi, N. et al "Production of L-phenylalanine by fermentation," Patent. JP 1993076352-A 2 03/30/93
E05108		Aspartokinase	Fugono, N et al. "Gene DNA coding Aspartokinase and its use," Patent: JP 1993184366-A 1 07/27/93
E05112		Dihydro-dipichorinate synthetase	Hatakeyama, K. et al "Gene DNA coding dihydrodipicolinic acid synthetase and its use," Patent: JP 1993184371-A 1 07/27/93

GenBank™ Accession No.	Gene Name	Gene Function	Reference
E05776		Diaminopimelic acid dehydrogenase	Kobayashi, M et al. "Gene DNA coding Diaminopimelic acid dehydrogenase and its use," Patent. JP 1993284970-A 1 11/02/93
E05779		Threonine synthase	Kohama, K. et al. "Gene DNA coding threonine synthase and its use," Patent. JP 1993284972-A 1 11/02/93
E06110		Prephenate dehydratase	Kikuchi, T. et al. "Production of L-phenylalanine by fermentation method," Patent. JP 1993344881-A 1 12/27/93
E06111		Mutated Prephenate dehydratase	Kikuchi, T. et al. "Production of L-phenylalanine by fermentation method," Patent: JP 1993344881-A 1 12/27/93
E06146		Acetohydroxy acid synthetase	Inui, M. et al. "Gene capable of coding Acetohydroxy acid synthetase and its use," Patent JP 1993344893-A 1 12/27/93
E06825		Aspartokinase	Sugimoto, M. et al "Mutant aspartokinase gene," patent JP 1994062866-A 1 03/08/94
E06826		Mutated aspartokinase alpha subunit	Sugimoto, M et al. "Mutant aspartokinase gene," patent: JP 1994062866-A 1 03/08/94
E06827		Mutated aspartokinase alpha subunit	Sugimoto, M. et al. "Mutant aspartokinase gene," patent JP 1994062866-A 1 03/08/94
E07701	secY		Honno, N. et al. "Gene DNA participating in integration of membraneous protein to membrane," Patent JP 1994169780-A 1 06/21/94
E08177		Aspartokinase	Sato, Y et al "Genetic DNA capable of coding Aspartokinase released from feedback inhibition and its utilization," Patent: JP 1994261766-A 1 09/20/94
E08178, E08179, E08180, E08181, E08182		Feedback inhibition-released Aspartokinase	Sato, Y. et al. "Genetic DNA capable of coding Aspartokinase released from feedback inhibition and its utilization," Patent: JP 1994261766-A 1 09/20/94
E08232		Acetohydroxy-acid isomeroreductase	Inui, M. et al. "Gene DNA coding acetohydroxy acid isomeroreductase," Patent: JP 1994277067-A 1 10/04/94
E08234	secE		Asai, Y. et al. "Gene DNA coding for translocation machinery of protein," Patent JP 1994277073-A 1 10/04/94
E08643		FT aminotransferase and desthiobiotin synthetase promoter region	Hatakeyama, K. et al. "DNA fragment having promoter function in coryneform bacterium," Patent JP 1995031476-A 1 02/03/95
E08646		Biotin synthetase	Hatakeyama, K. et al. "DNA fragment having promoter function in coryneform bacterium," Patent: JP 1995031476-A 1 02/03/95

GenBank™ Accession No.	Gene Name	Gene Function	Reference
E08649		Aspartase	Kohama, K. et al "DNA fragment having promoter function in coryneform bacterium," Patent. JP 1995031478-A 1 02/03/95
E08900		Dihydrodipicolinate reductase	Madori, M. et al "DNA fragment containing gene coding Dihydrodipicolinate acid reductase and utilization thereof," Patent: JP 1995075578-A 1 03/20/95
E08901		Diaminopimelic acid decarboxylase	Madori, M. et al. "DNA fragment containing gene coding Diaminopimelic acid decarboxylase and utilization thereof," Patent. JP 1995075579-A 1 03/20/95
E12594		Serine hydroxymethyltransferase	Hatakeyama, K. et al. "Production of L-trypophan," Patent· JP 1997028391-A 1 02/04/97
E12760, E12759, E12758		transposase	Moriya, M. et al "Amplification of gene using artificial transposon," Patent: JP 1997070291-A 03/18/97
E12764		Arginyl-tRNA synthetase; diaminopimelic acid decarboxylase	Moriya, M. et al. "Amplification of gene using artificial transposon," Patent. JP 1997070291-A 03/18/97
E12767		Dihydrodipicolinic acid synthetase	Moriya, M. et al. "Amplification of gene using artificial transposon," Patent: JP 1997070291-A 03/18/97
E12770		aspartokinase	Moriya, M. et al. "Amplification of gene using artificial transposon," Patent: JP 1997070291-A 03/18/97
E12773		Dihydrodipicolinic acid reductase	Moriya, M. et al. "Amplification of gene using artificial transposon," Patent· JP 1997070291-A 03/18/97
E13655		Glucose-6-phosphate dehydrogenase	Hatakeyama, K et al. "Glucose-6-phosphate dehydrogenase and DNA capable of coding the same," Patent· JP 1997224661-A 1 09/02/97
L01508	IlvA	Threonine dehydratase	Moeckel, B. et al. "Functional and structural analysis of the threonine dehydratase of Corynebacterium glutamicum," *J. Bacteriol*, 174.8065-8072 (1992)
L07603	EC 4.2 1.15	3-deoxy-D-arabinoheptulosonate-7-phosphate synthase	Chen, C et al. "The cloning and nucleotide sequence of Corynebacterium glutamicum 3-deoxy-D-arabinoheptulosonate-7-phosphate synthase gene," *FEMS Microbiol Lett.*, 107:223-230 (1993)
L09232	IlvB; ilvN; ilvC	Acetohydroxy acid synthase large subunit; Acetohydroxy acid synthase small subunit; Acetohydroxy acid isomeroreductase	Keilhauer, C. et al. "Isoleucine synthesis in Corynebacterium glutamicum: molecular analysis of the ilvB-ilvN-ilvC operon," *J. Bacteriol.*, 175(17).5595-5603 (1993)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
L18874	PtsM	Phosphoenolpyruvate sugar phosphotransferase	Fouet, A et al "Bacillus subtilis sucrose-specific enzyme II of the phosphotransferase system: expression in Escherichia coli and homology to enzymes II from enteric bacteria," *PNAS USA*, 84(24) 8773-8777 (1987); Lee, J.K. et al. "Nucleotide sequence of the gene encoding the Corynebacterium glutamicum mannose enzyme II and analyses of the deduced protein sequence," *FEMS Microbiol Lett*, 119(1-2).137-145 (1994)
L27123	aceB	Malate synthase	Lee, H-S. et al. "Molecular characterization of aceB, a gene encoding malate synthase in Corynebacterium glutamicum," *J Microbiol. Biotechnol*, 4(4) 256-263 (1994)
L27126		Pyruvate kinase	Jetten, M. S. et al. "Structural and functional analysis of pyruvate kinase from Corynebacterium glutamicum," *Appl. Environ Microbiol.*, 60(7):2501-2507 (1994)
L28760	aceA	Isocitrate lyase	
L35906	dtxI	Diphtheria toxin repressor	Oguiza, J.A. et al. "Molecular cloning, DNA sequence analysis, and characterization of the Corynebacterium diphtheriae dtxR from Brevibacterium lactofermentum," *J. Bacteriol*, 177(2):465-467 (1995)
M13774		Prephenate dehydratase	Follettie, M.T. et al. "Molecular cloning and nucleotide sequence of the Corynebacterium glutamicum pheA gene," *J Bacteriol.*, 167:695-702 (1986)
M16175	5S rRNA		Park, Y-H et al. "Phylogenetic analysis of the coryneform bacteria by 56 rRNA sequences," *J. Bacteriol*, 169:1801-1806 (1987)
M16663	trpE	Anthranilate synthase, 5' end	Sano, K. et al "Structure and function of the trp operon control regions of Brevibacterium lactofermentum, a glutamic-acid-producing bacterium," *Gene*, 52.191-200 (1987)
M16664	trpA	Tryptophan synthase, 3' end	Sano, K. et al. "Structure and function of the trp operon control regions of Brevibacterium lactofermentum, a glutamic-acid-producing bacterium," *Gene*, 52:191-200 (1987)
M25819		Phosphoenolpyruvate carboxylase	O'Regan, M. et al. "Cloning and nucleotide sequence of the Phosphoenolpyruvate carboxylase-coding gene of Corynebacterium glutamicum ATCC13032," *Gene*, 77(2):237-251 (1989)
M85106		23S rRNA gene insertion sequence	Roller, C. et al. "Gram-positive bacteria with a high DNA G+C content are characterized by a common insertion within their 23S rRNA genes," *J. Gen Microbiol*, 138.1167-1175 (1992)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
M85107, M85108		23S rRNA gene insertion sequence	Rollet, C. et al. "Gram-positive bacteria with a high DNA G+C content are characterized by a common insertion within their 23S rRNA genes," *J. Gen Microbiol.*, 138:1167-1175 (1992)
M89931	aecD; bmQ, yhbw	Beta C-S lyase; branched-chain amino acid uptake carrier, hypothetical protein yhbw	Rossol, I. et al "The Corynebacterium glutamicum aecD gene encodes a C-S lyase with alpha, beta-elimination activity that degrades aminoethylcysteine," *J. Bacteriol.*, 174(9):2968-2977 (1992); Tauch, A. et al. "Isoleucine uptake in Corynebacterium glutamicum ATCC 13032 is directed by the bmQ gene product," *Arch Microbiol.*, 169(4):303-312 (1998)
S59299	trp	Leader gene (promoter)	Henry, D.M. et al. "Cloning of the trp gene cluster from a tryptophan-hyperproducing strain of Corynebacterium glutamicum: identification of a mutation in the trp leader sequence," *Appl. Environ. Microbiol*, 59(3):791-799 (1993)
U11545	trpD	Anthranilate phosphoribosyltransferase	O'Gara, J.P. and Dunican, L.K (1994) Complete nucleotide sequence of the Corynebacterium glutamicum ATCC 21850 trpD gene." Thesis, Microbiology Department, University College Galway, Ireland.
U13922	cglIM; cglIR, clglIR	Putative type II 5-cytosoine methyltransferase; putative type II restriction endonuclease; putative type I or type III restriction endonuclease	Schafer, A. et al "Cloning and characterization of a DNA region encoding a stress-sensitive restriction system from Corynebacterium glutamicum ATCC 13032 and analysis of its role in intergeneric conjugation with Escherichia coli," *J Bacteriol.*, 176(23):7309-7319 (1994); Schafer, A. et al. "The Corynebacterium glutamicum cglIM gene encoding a 5-cytosine in an McrBC-deficient Escherichia coli strain," *Gene*, 203(2):95-101 (1997)
U14965	recA		
U31224	ppx		Ankri, S et al. "Mutations in the Corynebacterium glutamicumproline biosynthetic pathway: A natural bypass of the proA step," *J Bacteriol.*, 178(15):4412-4419 (1996)
U31225	proC	L-proline: NADP+ 5-oxidoreductase	Ankri, S. et al "Mutations in the Corynebacterium glutamicumproline biosynthetic pathway. A natural bypass of the proA step," *J Bacteriol*, 178(15):4412-4419 (1996)
U31230	obg; proB, unkdh	?;gamma glutamyl kinase;similar to D-isomer specific 2-hydroxyacid dehydrogenases	Ankri, S. et al. "Mutations in the Corynebacterium glutamicumproline biosynthetic pathway. A natural bypass of the proA step," *J. Bacteriol*, 178(15):4412-4419 (1996)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
U31281	bioB	Biotin synthase	Serebriiskii, I.G., "Two new members of the bio B superfamily: Cloning, sequencing and expression of bio B genes of Methylobacillus flagellatum and Corynebacterium glutamicum," *Gene*, 175:15-22 (1996)
U35023	thtR; accBC	Thiosulfate sulfurtransferase; acyl CoA carboxylase	Jager, W. et al. "A Corynebacterium glutamicum gene encoding a two-domain protein similar to biotin carboxylases and biotin-carboxyl-carrier proteins," *Arch Microbiol.*, 166(2);76-82 (1996)
U43535	cmr	Multidrug resistance protein	Jager, W. et al. "A Corynebacterium glutamicum gene conferring multidrug resistance in the heterologous host Escherichia coli," *J Bacteriol*, 179(7):2449-2451 (1997)
U43536	clpB	Heat shock ATP-binding protein	
U53587	aphA-3	3'5''-aminoglycoside phosphotransferase	
U89648		Corynebacterium glutamicum unidentified sequence involved in histidine biosynthesis, partial sequence	
X04960	trpA; trpB; trpC; trpD; trpE; trpG; trpL	Tryptophan operon	Matsui, K. et al. "Complete nucleotide and deduced amino acid sequences of the Brevibacterium lactofermentum tryptophan operon," *Nucleic Acids Res*, 14(24):10113-10114 (1986)
X07563	lys A	DAP decarboxylase (meso-diaminopimelate decarboxylase, EC 4.1.1.20)	Yeh, P. et al "Nucleic sequence of the lysA gene of Corynebacterium glutamicum and possible mechanisms for modulation of its expression," *Mol Gen. Genet*, 212(1):112-119 (1988)
X14234	EC 4 1.1.31	Phosphoenolpyruvate carboxylase	Eikmanns, B.J. et al. "The Phosphoenolpyruvate carboxylase gene of Corynebacterium glutamicum: Molecular cloning, nucleotide sequence, and expression," *Mol Gen. Genet*, 218(2):330-339 (1989); Lepiniec, L. et al. "Sorghum Phosphoenolpyruvate carboxylase gene family: structure, function and molecular evolution," *Plant Mol. Biol.*, 21 (3):487-502 (1993)
X17313	fda	Fructose-bisphosphate aldolase	Von der Osten, C.H. et al. "Molecular cloning, nucleotide sequence and fine-structural analysis of the Corynebacterium glutamicum fda gene: structural comparison of C glutamicum fructose-1, 6-biphosphate aldolase to class I and class II aldolases," *Mol. Microbiol*,
X53993	dapA	L-2, 3-dihydrodipicolinate synthetase (EC 4 2.1.52)	Bonnassie, S. et al. "Nucleic sequence of the dapA gene from Corynebacterium glutamicum," *Nucleic Acids Res*, 18(21):6421 (1990)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
X54223		AttB-related site	Cianciotto, N. et al. "DNA sequence homology between att B-related sites of Corynebacterium diphtheriae, Corynebacterium ulcerans, Corynebacterium glutamicum , and the attP site of lambdacorynephage," *FEMS Microbiol. Lett* , 66:299-302 (1990)
X54740	argS; lysA	Arginyl-tRNA synthetase; Diaminopimelate decarboxylase	Marcel, T. et al. "Nucleotide sequence and organization of the upstream region of the Corynebacterium glutamicum lysA gene," *Mol. Microbiol* , 4(11):1819-1830 (1990)
X55994	trpL; trpE	Putative leader peptide; anthranilate synthase component 1	Heery, D M. et al. "Nucleotide sequence of the Corynebacterium glutamicum trpE gene," *Nucleic Acids Res.*, 18(23):7138 (1990)
X56037	thrC	Threonine synthase	Han, K.S. et al. "The molecular structure of the Corynebacterium glutamicum threonine synthase gene," *Mol. Microbiol* , 4(10).1693-1702 (1990)
X56075	attB-related site	Attachment site	Cianciotto, N. et al. "DNA sequence homology between att B-related sites of Corynebacterium diphtheriae, Corynebacterium ulcerans, Corynebacterium glutamicum , and the attP site of lambdacorynephage," *FEMS Microbiol. Lett.*, 66:299-302 (1990)
X57226	lysC-alpha; lysC-beta; asd	Aspartokinase-alpha subunit, Aspartokinase-beta subunit; aspartate beta semialdehyde dehydrogenase	Kalinowski, J et al. "Genetic and biochemical analysis of the Aspartokinase from Corynebacterium glutamicum," *Mol Microbiol.*, 5(5).1197-1204 (1991); Kalinowski, J. et al "Aspartokinase genes lysC alpha and lysC beta overlap and are adjacent to the aspertate beta-semialdehyde dehydrogenase gene asd in Corynebacterium glutamicum," *Mol Gen Genet* , 224(3):317-324 (1990)
X59403	gap,pgk; tpi	Glyceraldehyde-3-phosphate; phosphoglycerate kinase, triosephosphate isomerase	Eikmanns, B.J "Identification, sequence analysis, and expression of a Corynebacterium glutamicum gene cluster encoding the three glycolytic enzymes glyceraldehyde-3-phosphate dehydrogenase, 3-phosphoglycerate kinase, and triosephosphate isomeras," *J. Bacteriol* , 174(19):6076-6086 (1992)
X59404	gdh	Glutamate dehydrogenase	Bormann, E.R et al. "Molecular analysis of the Corynebacterium glutamicum gdh gene encoding glutamate dehydrogenase," *Mol Microbiol.*, 6(3):317-326 (1992)
X60312	lysI	L-lysine permease	Seep-Feldhaus, A.H. et al. "Molecular analysis of the Corynebacterium glutamicum lysI gene involved in lysine uptake," *Mol Microbiol* , 5(12).2995-3005 (1991)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
X66078	copI	PsI protein	Joliff, G. et al "Cloning and nucleotide sequence of the cspl gene encoding PS1, one of the two major secreted proteins of Corynebacterium glutamicum: The deduced N-terminal region of PS1 is similar to the Mycobacterium antigen 85 complex," *Mol. Microbiol*, 6(16).2349-2362 (1992)
X66112	gli	Citrate synthase	Eikmanns, B.J. et al. "Cloning sequence, expression and transcriptional analysis of the Corynebacterium glutamicum gltA gene encoding citrate synthase," *Microbiol*, 140:1817-1828 (1994)
X67737	dapB	Dihydrodipicolinate reductase	
X69103	csp2	Surface layer protein PS2	Peyret, J.L. et al. "Characterization of the cspB gene encoding PS2, an ordered surface-layer protein in Corynebacterium glutamicum," *Mol Microbiol*, 9(1):97-109 (1993)
X69104		IS3 related insertion element	Bonamy, C et al. "Identification of IS1206, a Corynebacterium glutamicum IS3-related insertion sequence and phylogenetic analysis," *Mol. Microbiol*, 14(3):571-581 (1994)
X70959	leuA	Isopropylmalate synthase	Patek, M. et al. "Leucine synthesis in Corynebacterium glutamicum: enzyme activities, structure of leuA, and effect of leuA inactivation on lysine synthesis," *Appl Environ. Microbiol*, 60(1).133-140 (1994)
X71489	icd	Isocitrate dehydrogenase (NADP+)	Eikmanns, B.J. et al "Cloning sequence analysis, expression, and inactivation of the Corynebacterium glutamicum icd gene encoding isocitrate dehydrogenase and biochemical characterization of the enzyme," *J. Bacteriol.*, 177(3).774-782 (1995)
X72855	GDHA	Glutamate dehydrogenase (NADP+)	
X75083, X70584	mtrA	5-methyltryptophan resistance	Heery, D.M. et al. "A sequence from a tryptophan-hyperproducing strain of Corynebacterium glutamicum encoding resistance to 5-methyltryptophan," *Biochem Biophys Res. Commun*, 201(3):1255-1262 (1994)
X75085	recA		Fitzpatrick, R. et al. "Construction and characterization of recA mutant strains of Corynebacterium glutamicum and Brevibacterium lactofermentum," *Appl. Microbiol Biotechnol*, 42(4).575-580 (1994)
X75504	aceA; thiX	Partial Isocitrate lyase; ?	Reinscheid, D.J. et al. "Characterization of the isocitrate lyase gene from Corynebacterium glutamicum and biochemical analysis of the enzyme," *J. Bacteriol*, 176(12):3474-3483 (1994)
X76875		ATPase beta-subunit	Ludwig, W et al. "Phylogenetic relationships of bacteria based on comparative sequence analysis of elongation factor Tu and ATP-synthase beta-subunit genes," *Antonie Van Leeuwenhoek*, 64:285-305 (1993)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
X77034	tuf	Elongation factor Tu	Ludwig, W. et al. "Phylogenetic relationships of bacteria based on comparative sequence analysis of elongation factor Tu and ATP-synthase beta-subunit genes," *Antonie Van Leeuwenhoek*, 64:285-305 (1993)
X77384	recA		Billman-Jacobe, H. "Nucleotide sequence of a recA gene from Corynebacterium glutamicum," *DNA Seq.*, 4(6):403-404 (1994)
X78491	aceB	Malate synthase	Reinscheid, D.J. et al. "Malate synthase from Corynebacterium glutamicum pta-ack operon encoding phosphotransacetylase: sequence analysis," *Microbiology*, 140:3099-3108 (1994)
X80629	16S rDNA	16S ribosomal RNA	Rainey, F.A. et al. "Phylogenetic analysis of the genera Rhodococcus and Norcardia and evidence for the evolutionary origin of the genus Norcardia from within the radiation of Rhodococcus species," *Microbiol.*, 141:523-528 (1995)
X81191	gluA; gluB; gluC; gluD	Glutamate uptake system	Kronemeyer, W. et al "Structure of the gluABCD cluster encoding the glutamate uptake system of Corynebacterium glutamicum," *J Bacteriol*, 177(5):1152-1158 (1995)
X81379	dapE	Succinyldiaminopimelate desuccinylase	Wehrmann, A et al "Analysis of different DNA fragments of Corynebacterium glutamicum complementing dapE of Escherichia coli," *Microbiology*, 40:3349-56 (1994)
X82061	16S rDNA	16S ribosomal RNA	Ruimy, R et al. "Phylogeny of the genus Corynebacterium deduced from analyses of small-subunit ribosomal DNA sequences," *Int. J Syst Bacteriol*, 45(4):740-746 (1995)
X82928	asd; lysC	Aspartate-semialdehyde dehydrogenase; ?	Serebrijski, I. et al. "Multicopy suppression by asd gene and osmotic stress-dependent complementation by heterologous proA in proA mutants," *J. Bacteriol*, 177(24):7255-7260 (1995)
X82929	proA	Gamma-glutamyl phosphate reductase	Serebrijski, I. et al. "Multicopy suppression by asd gene and osmotic stress-dependent complementation by heterologous proA in proA mutants," *J. Bacteriol*, 177(24):7255-7260 (1995)
X84257	16S rDNA	16S ribosomal RNA	Pascual, C. et al. "Phylogenetic analysis of the genus Corynebacterium based on 16S rRNA gene sequences," *Int. J. Syst Bacteriol.*, 45(4):724-728 (1995)
X85965	aroP; dapE	Aromatic amino acid permease; ?	Wehrmann, A. et al. "Functional analysis of sequences adjacent to dapE of Corynebacterium glutamicumproline reveals the presence of aroP, which encodes the aromatic amino acid transporter," *J. Bacteriol*, 177(20):5991-5993 (1995)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
X86157	argB, argC; argD; argF; argJ	Acetylglutamate kinase; N-acetyl-gamma-glutamyl-phosphate reductase; acetylornithine aminotransferase; ornithine carbamoyltransferase; glutamate N-acetyltransferase	Sakanyan, V. et al. "Genes and enzymes of the acetyl cycle of arginine biosynthesis in Corynebacterium glutamicum: enzyme evolution in the early steps of the arginine pathway," *Microbiology*, 142:99-108 (1996)
X89084	pta; ackA	Phosphate acetyltransferase, acetate kinase	Reinscheid, D.J. et al "Cloning, sequence analysis, expression and inactivation of the Corynebacterium glutamicum pta-ack operon encoding phosphotransacetylase and acetate kinase," *Microbiology*, 145:503-513 (1999)
X89850	attB	Attachment site	Le Marrec, C et al. "Genetic characterization of site-specific integration functions of phi AAU2 infecting "Arthrobacter aureus C70," *J. Bacteriol*, 178(7):1996-2004 (1996)
X90356		Promoter fragment F1	Patek, M. et al. "Promoters from Corynebacterium glutamicum: cloning, molecular analysis and search for a consensus motif," *Microbiology*, 142:1297-1309 (1996)
X90357		Promoter fragment F2	Patek, M. et al. "Promoters from Corynebacterium glutamicum: cloning, molecular analysis and search for a consensus motif," *Microbiology*, 142:1297-1309 (1996)
X90358		Promoter fragment F10	Patek, M. et al. "Promoters from Corynebacterium glutamicum: cloning, molecular analysis and search for a consensus motif," *Microbiology*, 142:1297-1309 (1996)
X90359		Promoter fragment F13	Patek, M et al. "Promoters from Corynebacterium glutamicum: cloning, molecular analysis and search for a consensus motif," *Microbiology*, 142:1297-1309 (1996)
X90360		Promoter fragment F22	Patek, M. et al. "Promoters from Corynebacterium glutamicum: cloning, molecular analysis and search for a consensus motif," *Microbiology*, 142:1297-1309 (1996)
X90361		Promoter fragment F34	Patek, M. et al "Promoters from Corynebacterium glutamicum: cloning, molecular analysis and search for a consensus motif," *Microbiology*, 142:1297-1309 (1996)
X90362		Promoter fragment F37	Patek, M et al "Promoters from Corynebacterium glutamicum: cloning, molecular analysis and search for a consensus motif," *Microbiology*, 142:1297-1309 (1996)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
X90363		Promoter fragment F45	Patek, M et al "Promoters from Corynebacterium glutamicum: cloning, molecular analysis and search for a consensus motif," *Microbiology*, 142:1297-1309 (1996)
X90364		Promoter fragment F64	Patek, M. et al. "Promoters from Corynebacterium glutamicum: cloning, molecular analysis and search for a consensus motif," *Microbiology*, 142:1297-1309 (1996)
X90365		Promoter fragment F75	Patek, M. et al. "Promoters from Corynebacterium glutamicum: cloning, molecular analysis and search for a consensus motif," *Microbiology*, 142:1297-1309 (1996)
X90366		Promoter fragment PF101	Patek, M. et al. "Promoters from Corynebacterium glutamicum: cloning, molecular analysis and search for a consensus motif," *Microbiology*, 142:1297-1309 (1996)
X90367		Promoter fragment PF104	Patek, M. et al. "Promoters from Corynebacterium glutamicum: cloning, molecular analysis and search for a consensus motif," *Microbiology*, 142:1297-1309 (1996)
X90368		Promoter fragment PF109	Patek, M. et al. "Promoters from Corynebacterium glutamicum: cloning, molecular analysis and search for a consensus motif," *Microbiology*, 142:1297-1309 (1996)
X93513	amt	Ammonium transport system	Siewe, R.M. et al. "Functional and genetic characterization of the (methyl) ammonium uptake carrier of Corynebacterium glutamicum," *J Biol Chem*, 271(10):5398-5403 (1996)
X93514	betP	Glycine betaine transport system	Peter, H et al. "Isolation, characterization, and expression of the Corynebacterium glutamicum betP gene, encoding the transport system for the compatible solute glycine betaine," *J Bacteriol*, 178(17):5229-5234 (1996)
X95649	orf4		Patek, M. et al. "Identification and transcriptional analysis of the dapB-ORF2-dapA-ORF4 operon of Corynebacterium glutamicum, encoding two enzymes involved in L-lysine synthesis," *Biotechnol Lett*, 19:1113-1117 (1997)
X96471	lysE; lysG	Lysine exporter protein, Lysine export regulator protein	Vrljic, M. et al. "A new type of transporter with a new type of cellular function: L-lysine export from Corynebacterium glutamicum," *Mol Microbiol*, 22(5):815-826 (1996)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
X96580	panB; panC; xylB	3-methyl-2-oxobutanoate hydroxymethyltransferase; pantoate-beta-alanine ligase; xylulokinase	Sahm, H. et al. "D-pantothenate synthesis in Corynebacterium glutamicum and use of panBC and genes encoding L-valine synthesis for D-pantothenate overproduction," *Appl. Environ. Microbiol.*, 65(5):1973-1979 (1999)
X96962		Insertion sequence IS1207 and transposase	
X99289		Elongation factor P	Ramos, A. et al. "Cloning, sequencing and expression of the gene encoding elongation factor P in the amino-acid producer Brevibacterium lactofermentum (Corynebacterium glutamicum ATCC 13869)," *Gene*, 198:217-222 (1997)
Y00140	thrB	Homoserine kinase	Malcos, L.M. et al. "Nucleotide sequence of the homoserine kinase (thrB) gene of the Brevibacterium lactofermentum," *Nucleic Acids Res.*, 15(9):3922 (1987)
Y00151	ddh	Meso-diaminopimelate D-dehydrogenase (EC 1.4.1.16)	Ishino, S. et al. "Nucleotide sequence of the meso-diaminopimelate D-dehydrogenase gene from Corynebacterium glutamicum," *Nucleic Acids Res.*, 15(9):3917 (1987)
Y00476	thrA	Homoserine dehydrogenase	Mateos, L.M. et al. "Nucleotide sequence of the homoserine dehydrogenase (thrA) gene of the Brevibacterium lactofermentum," *Nucleic Acids Res.*, 15(24):10598 (1987)
Y00546	hom; thrB	Homoserine dehydrogenase; homoserine kinase	Peoples, O.P. et al. "Nucleotide sequence and fine structural analysis of the Corynebacterium glutamicum hom-thrB operon," *Mol. Microbiol.*, 2(1):63-72 (1988)
Y08964	murC, ftsQ/divD; ftsZ	UPD-N-acetylmuramate-alanine ligase, division initiation protein or cell division protein; cell division protein	Honrubia, M.P. et al. "Identification, characterization, and chromosomal organization of the ftsZ gene from Brevibacterium lactofermentum," *Mol. Gen. Genet.*, 259(1):97-104 (1998)
Y09163	putP	High affinity proline transport system	Peter, H. et al. "Isolation of the putP gene of Corynebacterium glutamicumproline and characterization of a low-affinity uptake system for compatible solutes," *Arch. Microbiol.*, 168(2):143-151 (1997)
Y09548	pyc	Pyruvate carboxylase	Peters-Wendisch, P.G. et al. "Pyruvate carboxylase from Corynebacterium glutamicum: characterization, expression and inactivation of the pyc gene," *Microbiology*, 144:915-927 (1998)
Y09578	leuB	3-isopropylmalate dehydrogenase	Patek, M. et al. "Analysis of the leuB gene from Corynebacterium glutamicum," *Appl. Microbiol. Biotechnol.*, 50(1):42-47 (1998)
Y12472		Attachment site bacteriophage Phi-16	Moreau, S. et al. "Site-specific integration of corynephage Phi-16: The construction of an integration vector," *Microbiol.*, 145:539-548 (1999)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
Y12537	proP	Proline/ectoine uptake system protein	Peter, H. et al. "Corynebacterium glutamicum is equipped with four secondary carriers for compatible solutes. Identification, sequencing, and characterization of the proline/ectoine uptake system, ProP, and the ectoine/proline/glycine betaine carrier, EctP," *J. Bacteriol.*, 180(22):6005-6012 (1998)
Y13221	glnA	Glutamine synthetase I	Jakoby, M. et al. "Isolation of Corynebacterium glutamicum glnA gene encoding glutamine synthetase I," *FEMS Microbiol Lett.*, 154(1):81-88 (1997)
Y16642	lpd	Dihydrolipoamide dehydrogenase	
Y18059		Attachment site Corynephage 304L	Moreau, S. et al "Analysis of the integration functions of φ304L. An integrase module among corynephages," *Virology*, 255(1):150-159 (1999)
Z21501	argS; lysA	Arginyl-tRNA synthetase; diaminopimelate decarboxylase (partial)	Oguiza, J.A. et al. "A gene encoding arginyl-tRNA synthetase is located in the upstream region of the lysA gene in Brevibacterium lactofermentum. Regulation of argS-lysA cluster expression by arginine," *J. Bacteriol.*, 175(22):7356-7362 (1993)
Z21502	dapA; dapB	Dihydrodipicolinate synthase; dihydrodipicolinate reductase	Pisabarro, A. et al. "A cluster of three genes (dapA, orf2, and dapB) of Brevibacterium lactofermentum encodes dihydrodipicolinate reductase, and a third polypeptide of unknown function," *J. Bacteriol.*, 175(9):2743-2749 (1993)
Z29563	thrC	Threonine synthase	Malumbres, M. et al. "Analysis and expression of the thrC gene of the encoded threonine synthase," *Appl. Environ. Microbiol.*, 60(7)2209-2219 (1994)
Z46753	16S rDNA	Gene for 16S ribosomal RNA	
Z49822	sigA	SigA sigma factor	Oguiza, J.A. et al "Multiple sigma factor genes in Brevibacterium lactofermentum: Characterization of sigA and sigB," *J. Bacteriol.*, 178(2):550-553 (1996)
Z49823	galE; dtxR	Catalytic activity UDP-galactose 4-epimerase; diphtheria toxin regulatory protein	Oguiza, J.A. et al "The galE gene encoding the UDP-galactose 4-epimerase of Brevibacterium lactofermentum is coupled transcriptionally to the dtxR gene," *Gene*, 177:103-107 (1996)
Z49824	orf1; sigB	?; SigB sigma factor	Oguiza, J.A. et al "Multiple sigma factor genes in Brevibacterium lactofermentum: Characterization of sigA and sigB," *J. Bacteriol.*, 178(2):550-553 (1996)
Z66534		Transposase	Correia, A. et al. "Cloning and characterization of an IS-like element present in the genome of Brevibacterium lactofermentum ATCC 13869," *Gene*, 170(1):91-94 (1996)

[1] A sequence for this gene was published in the indicated reference. However, the sequence obtained by the inventors of the present application is significantly longer than the published version. It is believed that the published version relied on an incorrect start codon, and thus represents only a fragment of the actual coding region.

TABLE 3: Corynebacterium and Brevibacterium Strains Which May be Used in the Practice of the Invention

Genus	species	ATCC	FERM	NRRL	CECT	NCIMB	CBS	NCTC	DSMZ
Brevibacterium	ammoniagenes	21054							
Brevibacterium	ammoniagenes	19350							
Brevibacterium	ammoniagenes	19351							
Brevibacterium	ammoniagenes	19352							
Brevibacterium	ammoniagenes	19353							
Brevibacterium	ammoniagenes	19354							
Brevibacterium	ammoniagenes	19355							
Brevibacterium	ammoniagenes	19356							
Brevibacterium	ammoniagenes	21055							
Brevibacterium	ammoniagenes	21077							
Brevibacterium	ammoniagenes	21553							
Brevibacterium	ammoniagenes	21580							
Brevibacterium	ammoniagenes	39101							
Brevibacterium	butanicum	21196							
Brevibacterium	divaricatum	21792	P928						
Brevibacterium	flavum	21474							
Brevibacterium	flavum	21129							
Brevibacterium	flavum	21518							
Brevibacterium	flavum			B11474					
Brevibacterium	flavum			B11472					
Brevibacterium	flavum	21127							
Brevibacterium	flavum	21128							
Brevibacterium	flavum	21427							
Brevibacterium	flavum	21475							
Brevibacterium	flavum	21517							
Brevibacterium	flavum	21528							
Brevibacterium	flavum	21529							
Brevibacterium	flavum			B11477					

Brevibacterium	flavum			B11478					
Brevibacterium	flavum	21127							
Brevibacterium	flavum			B11474					
Brevibacterium	healii	15527							
Brevibacterium	ketoglutamicum	21004							
Brevibacterium	ketoglutamicum	21089							
Brevibacterium	ketosoreductum	21914							
Brevibacterium	lactofermentum				70				
Brevibacterium	lactofermentum				74				
Brevibacterium	lactofermentum				77				
Brevibacterium	lactofermentum	21798							
Brevibacterium	lactofermentum	21799							
Brevibacterium	lactofermentum	21800							
Brevibacterium	lactofermentum	21801							
Brevibacterium	lactofermentum			B11470					
Brevibacterium	lactofermentum			B11471					
Brevibacterium	lactofermentum	21086							
Brevibacterium	lactofermentum	21420							
Brevibacterium	lactofermentum	21086							
Brevibacterium	lactofermentum	31269							
Brevibacterium	linens	9174							
Brevibacterium	linens	19391							
Brevibacterium	linens	8377							
Brevibacterium	paraffinolyticum					11160			
Brevibacterium	spec.						717.73		
Brevibacterium	spec.						717.73		
Brevibacterium	spec.	14604							
Brevibacterium	spec.	21860							
Brevibacterium	spec.	21864							
Brevibacterium	spec.	21865							
Brevibacterium	spec.	21866							
Brevibacterium	spec.	19240							

Corynebacterium	acetoacidophilum	21476							
Corynebacterium	acetoacidophilum	13870							
Corynebacterium	acetoglutamicum			B11473					
Corynebacterium	acetoglutamicum			B11475					
Corynebacterium	acetoglutamicum	15806							
Corynebacterium	acetoglutamicum	21491							
Corynebacterium	acetoglutamicum	31270							
Corynebacterium	acetophilum			B3671					
Corynebacterium	ammoniagenes	6872						2399	
Corynebacterium	ammoniagenes	15511							
Corynebacterium	fujiokense	21496							
Corynebacterium	glutamicum	14067							
Corynebacterium	glutamicum	39137							
Corynebacterium	glutamicum	21254							
Corynebacterium	glutamicum	21255							
Corynebacterium	glutamicum	31830							
Corynebacterium	glutamicum	13032							
Corynebacterium	glutamicum	14305							
Corynebacterium	glutamicum	15455							
Corynebacterium	glutamicum	13058							
Corynebacterium	glutamicum	13059							
Corynebacterium	glutamicum	13060							
Corynebacterium	glutamicum	21492							
Corynebacterium	glutamicum	21513							
Corynebacterium	glutamicum	21526							
Corynebacterium	glutamicum	21543							
Corynebacterium	glutamicum	13287							
Corynebacterium	glutamicum	21851							
Corynebacterium	glutamicum	21253							
Corynebacterium	glutamicum	21514							
Corynebacterium	glutamicum	21516							
Corynebacterium	glutamicum	21299							

Corynebacterium	glutamicum	21300							
Corynebacterium	glutamicum	39684							
Corynebacterium	glutamicum	21488							
Corynebacterium	glutamicum	21649							
Corynebacterium	glutamicum	21650							
Corynebacterium	glutamicum	19223							
Corynebacterium	glutamicum	13869							
Corynebacterium	glutamicum	21157							
Corynebacterium	glutamicum	21158							
Corynebacterium	glutamicum	21159							
Corynebacterium	glutamicum	21355							
Corynebacterium	glutamicum	31808							
Corynebacterium	glutamicum	21674							
Corynebacterium	glutamicum	21562							
Corynebacterium	glutamicum	21563							
Corynebacterium	glutamicum	21564							
Corynebacterium	glutamicum	21565							
Corynebacterium	glutamicum	21566							
Corynebacterium	glutamicum	21567							
Corynebacterium	glutamicum	21568							
Corynebacterium	glutamicum	21569							
Corynebacterium	glutamicum	21570							
Corynebacterium	glutamicum	21571							
Corynebacterium	glutamicum	21572							
Corynebacterium	glutamicum	21573							
Corynebacterium	glutamicum	21579							
Corynebacterium	glutamicum	19049							
Corynebacterium	glutamicum	19050							
Corynebacterium	glutamicum	19051							
Corynebacterium	glutamicum	19052							
Corynebacterium	glutamicum	19053							
Corynebacterium	glutamicum	19054							

Corynebacterium	glutamicum	19055							
Corynebacterium	glutamicum	19056							
Corynebacterium	glutamicum	19057							
Corynebacterium	glutamicum	19058							
Corynebacterium	glutamicum	19059							
Corynebacterium	glutamicum	19060							
Corynebacterium	glutamicum	19185							
Corynebacterium	glutamicum	13286							
Corynebacterium	glutamicum	21515							
Corynebacterium	glutamicum	21527							
Corynebacterium	glutamicum	21544							
Corynebacterium	glutamicum	21492							
Corynebacterium	glutamicum			B8183					
Corynebacterium	glutamicum			B8182					
Corynebacterium	glutamicum			B12416					
Corynebacterium	glutamicum			B12417					
Corynebacterium	glutamicum			B12418					
Corynebacterium	glutamicum			B11476					
Corynebacterium	glutamicum	21608							
Corynebacterium	lilium		P973						
Corynebacterium	nitrilophilus	21419				11594			
Corynebacterium	spec.		P4445						
Corynebacterium	spec.		P4446						
Corynebacterium	spec.	31088							
Corynebacterium	spec.	31089							
Corynebacterium	spec.	31090							
Corynebacterium	spec.	31090							
Corynebacterium	spec	31090							20145
Corynebacterium	spec.	15954							
Corynebacterium	spec.	21857							
Corynebacterium	spec.	21862							
Corynebacterium	spec.	21863							

ATCC: American Type Culture Collection, Rockville, MD, USA

FERM: Fermentation Research Institute, Chiba, Japan

NRRL: ARS Culture Collection, Northern Regional Research Laboratory, Peoria, IL, USA

CECT: Coleccion Espanola de Cultivos Tipo, Valencia, Spain

NCIMB: National Collection of Industrial and Marine Bacteria Ltd., Aberdeen, UK

CBS: Centraalbureau voor Schimmelcultures, Baarn, NL

NCTC: National Collection of Type Cultures, London, UK

DSMZ: Deutsche Sammlung von Mikroorganismen und Zellkulturen, Braunschweig, Germany

For reference see Sugawara, H et al. (1993) World directory of collections of cultures of microorganisms: Bacteria, fungi and yeasts (4th edn), World federation for culture collections world data center on microorganisms, Saimata, Japen.

>>RXA02747-amino acid sequence
(1-2076, translated) 692 residues
MNNPAQLRQD TEKEVLALLG SLVLPAGTAL AATGSLARSE LTPYSDLDLI LIHPPGATPD GVEDLWYPIW
DAKKRLDYSV RTPDECVAMI SADSTAALAM LDLRFVAGDE DLCAKTRRRI VEKWRQELNK NFDAVVDTAI
ARWRRSGPVV AMTRPDLKHG RGGLRDFELI KALALGHLCN LPQLDAQHQL LLDARTLLHV HARRSRDVLD
PEFAVDVAMD LGFVDRYHLG REIADAARAI DDGLTTALAT ARGILPRRTG FAFRNASRRP LDLDVVDANG
TIELSKKPDL NDPALPLRVA AAAATTGLPV AESTWVRLNE CPPLPEPWPA NAAGDFFRIL SSPKNSRRVV
KNMDRHGLWS RFVPEWDRIK GLMPREPSHI STIDEHSLNT VAGCALETVT VARPDLLVLG ALYHDIGKGF
PRPHEQVGAE MVARAASRMG LNLRDRASVQ TLVAEHTAVA KIAARLDPSS EGAVDKLLDA VRYDLVTLNL
LEVLTEADAK ATGPGVWTAR LEHALRIVCK RARDRLTDIR PVAPMIAPRS EIGLVERDGV FTVQWHGEDL
HRILGVIYAK GWTITAARML ANGQWSAEFD VRANGPQDFD PQHFLQAYQS GVFSEVPIPA LGITATFWHG
NTLEVRTELR TGAIFALLRT LPDALWINAV TRGATLIIQA ALKPGFDRAT VERSVVRSLA GS
>RXA02747-nucleotide sequence A: upstream
CACCGGCAAAGTCGGCGACGGCAAAGTGTGGATGACTAACATCGAAGAGCTGGTTCGTGTTCGTACCGGTGAGCGCG
GCGAAGCAGCCCTTTAAAAACTT
>RXA02747-nucleotide sequence B: coding region
ATGAATAATCCAGCCCAGCTGCGCCAAGATACTGAAAAGGAAGTCCTGGCGTTGCTGGGCTCTTTGGTTTTACCCGC
CGGCACCGCGCTTGCCGCCACCGGATCTTTGGCCAGGTCCGAACTCACGCCGTATTCCGATTTGGACCTCATTTTGA
TCCATCCACCAGGAGCCACCCCGGATGGCGTGGAGGATTTGTGGTACCCGATTTGGGACGCAAAAAAGCGTCTCGAC
TACTCCGTGCGCACCCCAGATGAGTGTGTGGCTATGATTTCTGCGGATTCCACTGCAGCCCTTGCCATGCTTGACCT
GCGGTTTGTCGCTGGCGATGAGGATCTGTGTGCCAAAACGCGCCGGCGCATCGTGGAGAAGTGGCGCCAGGAACTCA
ACAAAAACTTCGATGCCGTTGTGGACACCGCGATTGCCCGTTGGCGCCGCTCCGGACCCGTCGTGGCAATGACGCGG
CCAGATCTTAAACACGGCAGGGGAGGGCTGCGCGATTTCGAACTGATCAAGGCCCTCGCGCTCGGCCACCTATGCAA
CCTTCCACAGCTTGATGCGCAACACCAGCTGCTTCTCGACGCCCGCACCTTGCTGCACGTCCACGCGCGACGCTCCC
GCGACGTCCTTGACCCCGAATTTGCGGTGGATGTGGCCATGGATTTGGGCTTTGTTGACCGCTATCACCTGGGCCGG
GAGATCGCCGATGCAGCCCGCGCCATTGATGATGGCCTGACCACCGCGCTGGCCACCGCCCGTGGCATTTTGCCACG
TCGCACAGGTTTTGCATTCAGGAATGCTTCTCGACGCCCACTTGATCTTGATGTCGTCGACGCCAACGGCACCATCG
AATTGTCCAAAAAACCAGATCTTAATGATCCCGCACTTCCACTTCGAGTGGCCGCAGCCGCAGCAACCACCGGACTT
CCGGTGGCAGAATCAACCTGGGTTCGACTTAATGAATGCCCGCCACTTCCTGAGCCATGGCCTGCCAATGCAGCAGG
GGACTTCTTTCGGATTCTCTCCAGTCCGAAAAACTCACGCCGAGTGGTGAAAAATATGGATCGCCACGGATTGTGGT
CGCGTTTTGTTCCAGAATGGGACCGCATCAAAGGGCTTATGCCCCGTGAACCCAGCCATATTTCCACCATCGATGAA
CATAGTCTGAACACTGTTGCAGGATGTGCGCTAGAAACTGTGACCGTCGCGCGCCCCGATCTTTTAGTTTTGGGAGC
CTTGTACCACGACATTGGCAAGGGCTTCCCGCGTCCACACGAACAAGTAGGTGCAGAGATGGTGGCGAGGGCTGCAA
GCCGCATGGGATTGAACCTTCGCGATCGTGCCAGCGTGCAAACGCTGGTCGCCGAGCACACCGCGGTGGCCAAAATC
GCCGCGCGCCTTGATCCCTCCTCGGAGGGCGCCGTCGATAAGCTGCTTGATGCTGTTAGGTATGACCTGGTGACATT
GAATCTGCTTGAGGTGCTAACAGAAGCTGATGCGAAAGCCACGGGGCCTGGCGTGTGGACGGCGCGTTTGGAGCATG
CGCTGCGGATTGTGTGCAAGCGTGCGCGTGATCGCCTCACCGATATTCGCCCGGTTGCGCCGATGATTGCGCCACGT
AGTGAAATTGGTTTGGTGGAACGCGATGGCGTGTTCACAGTGCAATGGCACGGCGAAGACTTACATCGGATTCTTGG
CGTAATTTATGCCAAAGGATGGACAATCACCGCGGCGCGCATGCTGGCCAATGGTCAATGGAGTGCGGAATTTGATG
TCCGCGCAAACGGCCCCCAAGATTTTGATCCGCAGCATTTCCTGCAGGCATATCAATCCGGTGTGTTTTCCGAGGTT
CCCATTCCAGCACTTGGGATAACAGCCACATTTTGGCACGGGAACACTTTAGAAGTGCGCACTGAGCTTCGCACAGG
AGCTATTTTTGCCCTGCTCAGAACATTGCCCGATGCCCTCTGGATCAACGCTGTGACCCGCGGTGCGACCCTGATTA
TCCAGGCAGCACTGAAGCCCGGCTTCGATCGAGCAACGGTGGAACGCTCCGTAGTCAGGTCGTTGGCAGGTAGC
>RXA02747-nucleotide sequence C: downstream
TGACGTGACCTGAGCGGGGGCAA

>>RXA02699-amino acid sequence
(1-2148, translated) 716 residues

MSTVYRCLDL RLGRSMALKV MEEDFVDDPI FRQRSRREAR SMAQLNHPNL VNVYDFSATD GLVYLVMELI
TGGTLRELLA ERGPMPPHAA VGVMRGVLTG LAAAHRAGMV HRDIKPDNVL INSDHQVKLS DFGLVRAAHA
GQSQDNQIVG TVAYLSPEQV EGGEIGPASD VYSAGIVLFE LLTGTTPFSG EDDLDHAYAR LTEVVPAPSS
LIDGVPSLID ELVATATSIN PEDRFDDSGE FLSALEDVAT ELSLPAFRVP VPVNSAANRA NAQVPDAQPT
DMFTTHIPKT PEPDHTAIIP VASANETSIL PAQNMAQNMA QNPLQPPEPD FAPEPPPDTA LNIQDQELAR
ADEPEINTVS NRSKLKLTLW SIFVVAVIAA VAVGGWWFGS GRYGEIPQVL GMDEVQAVAV VEEAGFVAVA
EPQYDNEVPT GSIIGTEPSF GERLPRGEDV SVLVSQGRPV VPDLSEDRSL STVREELEQR TFVWVDGPGE
YSDDVPEGQV VSFTPSSGTQ LDVGETVQIH LSRGPAPVEI PDVSGMGVDQ ATRVLERAGL SVERTEEGFD
AETPNGDVYG TSPKVSTEVK RGTSVVLQVS NAISVPDVVG MTKDEATAAL AEEGLVVAST SIIPGEAASS
ADAVVTVEPE SGSRVDPAHP QVSLGLAGEI QVPSVVGRKV SDARSILEEA GLTLTTDADD NDRIYSQTPR
ARSEVSVGGE VTVRAF

>RXA02699-nucleotide sequence A: upstream

TTGTGTACCTTCCGACATACTGGAACGCATGGCAAACTTGAAGGTCGGTGACGTTTTAGAGGACAGGTATCGGATTG
AAACTCCGATTGCCCGGGGTGGT

>RXA02699-nucleotide sequence B: coding region

ATGTCTACCGTGTACAGGTGCCTTGATCTTCGTTTAGGACGTTCCATGGCGCTTAAAGTCATGGAAGAAGATTTCGT
TGATGATCCCATTTTCCGGCAGCGTTCCCGTAGGGAAGCTCGGTCAATGGCGCAGCTAAATCATCCAAATTTGGTCA
ATGTGTATGATTTTTCCGCTACTGACGGTTTGGTGTATCTGGTGATGGAGTTAATCACTGGTGGCACCTTGCGTGAG
TTGCTGGCTGAGCGGGGACCTATGCCCCCGCATGCTGCTGTGGGCGTTATGCGTGGGGTGCTCACGGGTCTCGCGGC
TGCCCACCGGGCGGGCATGGTGCACCGGGATATCAAGCCTGACAACGTGTTGATCAATAGTGATCACCAGGTGAAAC
TGTCTGATTTCGGCTTGGTTCGAGCGGCTCACGCCGGCCAGTCTCAGGACAATCAGATTGTGGGCACGGTGGCTTAT
CTTTCCCCTGAGCAGGTTGAGGGCGGTGAGATCGGGCCGGCCAGCGACGTGTATTCGGCAGGCATTGTGCTCTTTGA
GCTGCTCACAGGCACCACGCCTTTTTCGGGCGAGGATGATCTCGACCATGCATACGCCCGCCTTACGGAAGTCGTGC
CGGCACCGAGTTCGCTTATCGACGGCGTCCCCTCCCTCATCGATGAGCTTGTCGCGACAGCTACCTCCATTAATCCT
GAGGATCGTTTCGATGATTCTGGAGAGTTTTTGTCCGCACTGGAAGATGTCGCAACAGAGTTGAGCTTGCCGGCTTT
CCGGGTCCCTGTGCCGGTTAATTCCGCAGCCAATAGGGCTAATGCCCAGGTCCCGGATGCTCAGCCAACTGATATGT
TTACCACCCATATCCCCAAGACTCCTGAGCCTGATCACACTGCGATCATTCCGGTGGCCTCAGCAAATGAGACGTCG
ATTCTGCCTGCGCAAAACATGGCACAAAATATGGCGCAGAATCCGCTGCAACCTCCGGAACCTGATTTCGCCCCGGA
GCCACCTCCGGACACAGCGCTGAATATTCAAGATCAAGAGCTTGCGCGCGCCGATGAGCCAGAAATTAATACCGTCA
GCAATCGTTCCAAATTGAAGCTGACGTTGTGGTCAATTTTCGTGGTCGCAGTGATCGCTGCTGTTGCTGTTGGCGGT
TGGTGGTTCGGTTCAGGCCGTTACGGTGAGATTCCGCAGGTGTTGGGCATGGATGAGGTCCAGGCAGTAGCTGTTGT
AGAGGAAGCTGGTTTCGTGGCAGTGGCTGAACCTCAGTATGACAATGAGGTTCCCACTGGTTCGATTATTGGGACTG
AACCTTCTTTTGGTGAGCGCCTTCCTCGCGGCGAGGATGTTTCTGTCCTCGTCTCTCAAGGGCGTCCCGTGGTGCCG
GATCTTAGCGAGGATCGATCCTTAAGCACCGTTCGTGAAGAGTTGGAACAGCGCACGTTCGTCTGGGTTGATGGCCC
AGGTGAATATTCTGACGATGTTCCAGAAGGACAAGTAGTTTCTTTTACACCGTCGTCAGGCACGCAGCTTGATGTTG
GTGAAACCGTGCAGATCCATTTGAGCCGAGGCCCCGCCCCGGTTGAGATTCCTGATGTCTCTGGCATGGGAGTGGAT
CAGGCAACACGTGTGTTGGAGCGCGCAGGTTTGAGCGTCGAGCGTACTGAAGAAGGCTTTGATGCTGAGACACCAAA
TGGTGATGTCTACGGGACTTCGCCCAAGGTATCTACTGAGGTCAAGCGCGGAACCTCTGTTGTGCTGCAGGTGTCCA
ATGCTATTTCGGTACCGGATGTGGTGGGTATGACCAAGGACGAAGCCACCGCGGCGCTTGCGGAAGAAGGATTGGTC
GTGGCGTCGACAAGCATTATTCCTGGTGAGGCGGCGAGCTCCGCTGACGCCGTCGTGACCGTCGAGCCTGAATCCGG
CAGCCGCGTTGATCCAGCGCATCCGCAGGTCAGCCTCGGGTTAGCTGGGGAGATTCAAGTTCCAAGCGTGGTTGGAC
GTAAGGTTAGCGATGCTCGAAGCATTCTGGAAGAAGCCGGTTTAACGCTGACAACTGATGCGGACGACAACGATCGA
ATTTATAGTCAAACCCCTCGTGCACGCAGCGAAGTCTCGGTAGGGGGAGAAGTTACAGTAAGGGCGTTT

>RXA02699-nucleotide sequence C: downstream

TAGTGGTTCCCTCGTTGCAGCAA

>>RXA01272-amino acid sequence
(1-603, translated) 201 residues
MSNSFTILTV CTGNICRSPL AKQLLELELP GADIIRVDSA GVQAMVDSPM PEQSLEIARK QGIENPEEHR
AKQITEELVN QSDLILAMDR GHRKSIVQLS PRATRKVFTV VDLARLIEAT TDADLQEELN LAGDSVIDRL
HATVEAARLS RSELNPLDNL ADEDIVDPYG KSQSVYEASA SQLIPAIRLI ASYLNKALES A
>RXA01272-nucleotide sequence A: upstream
TATGGCTACGGAAATTACGGCTACGGCGACACCTCCAAAATCAATGCCCCTAAGCCCGACAACACCGAACTAACCAC
CACCGATGCTTCCAAGGCCAACA
>RXA01272-nucleotide sequence B: coding region
ATGAGCAATAGCTTCACTATTCTCACTGTCTGTACTGGAAACATTTGCCGCTCCCCGTTAGCTAAGCAGCTACTTGA
ACTTGAGCTTCCGGGGGCAGATATAATCCGCGTTGATTCCGCCGGTGTTCAGGCGATGGTTGATTCGCCTATGCCGG
AGCAATCTTTAGAAATCGCACGTAAACAGGGCATAGAAAACCCTGAGGAGCACCGAGCTAAGCAGATTACTGAGGAG
CTTGTAAACCAATCTGATCTGATTCTTGCGATGGATCGGGGGCATCGAAAATCCATTGTCCAGCTAAGCCCGCGTGC
AACCCGTAAGGTTTTCACTGTTGTTGATCTTGCCAGGTTAATTGAGGCAACAACTGATGCTGATCTGCAGGAAGAGC
TCAATCTGGCAGGGGATTCCGTGATCGATAGGCTGCATGCGACAGTTGAGGCTGCTCGTCTTAGCCGCAGTGAATTG
AATCCTCTGGATAACCTCGCAGATGAAGATATTGTTGACCCGTACGGAAAGAGTCAATCGGTTTATGAGGCATCGGC
GAGTCAGCTAATTCCAGCTATTCGTTTGATTGCTTCTTATTTGAACAAAGCACTGGAGTCTGCG
>RXA01272-nucleotide sequence C: downstream
TAATGGCGAGGAAGTATCGGGTG

>>RXA01826-amino acid sequence
(1-1938, translated) 646 residues
VTFVIADRYE LDAVIGSGGM SEVFAATDTL IGREVAVKML RIDLAKDPNF RERFRREAQN SGRLSHSSIV
AVFDTGEVDK DGTSVPYIVM ERVQGRNLRE VVTEDGVFTP VEAANILIPV CEALQASHDA GIIHRDVKPA
NIMITNTGGV KVMDFGIARA VNDSTSAMTQ TSAVIGTAQY LSPEQARGKP ADARSDIYAT GCVMYELVTG
KPPFEGESPF AVAYQHVQED PTPPSDFIAD LTPTSAVNVD AVVLTAMAKH PADRYQTASE MAADLGRLSR
NAVSHAARAH VETEETPEEP ETRFSTRTST QVAPAAGVAA ASTGSGSSSR KRGSRGLTAL AIVLSLGVVG
VAGAFTYDYF ANSSSTATSA IPNVEGLPQQ EALTELQAAG FVVNIVEEAS ADVAEGLVIR ANPSVGSEIR
QGATVTITVS TGREMINIPD VSGMTLEDAA RALEDVGLIL NQNVREETSD DVESGLVIDQ NPEAGQEVVV
GSSVSLTMSS GTESIRVPNL TGMNWSQAEQ NLISMGFNPT ASYLDSSEPE GEVLSVSSQG TELPKGSSIT
VEVSNGMLIQ APDLARMSTE QAISALRAAG WTAPDQSLIV GDPIHTAALV DQNKIGFQSP TPATLFRKDA
QVQVRLFEFD LAALVQ
>RXA01826-nucleotide sequence A: upstream
CTACACCTAGGACGAGTGCTAGCGTTCCAGTTGAGACTAATGCACCGGCTGATGATTTAATCGACGCCGTAAATGGC
CTATTGGATGTAGGAGGAGCGCA
>RXA01826-nucleotide sequence B: coding region
GTGACCTTCGTGATCGCTGATCGCTATGAACTGGATGCCGTCATCGGCTCCGGTGGCATGAGCGAGGTGTTCGCGGC
CACCGACACGCTCATTGGTCGGGAGGTCGCGGTAAAGATGCTGCGCATCGACCTTGCGAAAGATCCCAATTTCCGAG
AACGCTTCCGCAGGGAAGCCCAAAACTCCGGAAGGTTGAGCCACTCTTCGATCGTCGCTGTTTTTGACACCGGCGAA
GTAGACAAAGACGGCACCTCTGTTCCCTACATTGTGATGGAACGCGTGCAGGGTCGAAACCTGCGCGAAGTTGTCAC
CGAAGACGGCGTATTCACCCCAGTTGAGGCAGCCAACATCCTCATCCCTGTGTGTGAAGCGCTGCAGGCATCCCATG
ACGCCGGCATTATTCACCGCGATGTGAAACCCGCCAACATCATGATCACCAACACCGGTGGCGTGAAAGTCATGGAC
TTCGGCATCGCCCGCGCGGTCAACGATTCCACCTCCGCCATGACTCAAACCTCCGCAGTCATCGGCACCGCCCAGTA
CCTCTCCCCTGAGCAGGCCCGCGGCAAACCCGCCGATGCGCGTTCCGATATTTACGCCACCGGCTGCGTCATGTACG
AATTAGTCACCGGTAAGCCACCTTTTGAAGGCGAGTCCCCTTTCGCCGTGGCCTACCAACACGTCCAGGAAGACCCC
ACCCCTCCTTCGGATTTCATCGCGGACCTCACCCCGACCTCTGCTGTCAACGTGGATGCCGTGGTACTCACCGCCAT
GGCAAAACACCCCGCCGACCGCTACCAAACAGCCTCCGAAATGGCCGCTGACCTGGGCCGGCTATCCCGCAATGCAG
TCTCCCATGCCGCACGCGCGCATGTAGAAACAGAAGAAACCCCAGAAGAGCCCGAAACTCGCTTCTCGACGCGCACC
TCCACCCAAGTGGCCCCCGCCGCAGGCGTGGCTGCGGCCAGTACGGGGTCAGGGTCTTCTTCGCGTAAACGTGGATC
CAGAGGCCTCACCGCCCTGGCCATCGTGTTATCCCTAGGTGTCGTCGGCGTTGCCGGTGCCTTCACCTACGACTACT
TTGCCAACAGCTCCTCCACTGCAACCAGCGCGATCCCCAATGTGGAAGGCCTCCCGCAGCAAGAAGCTCTCACAGAA
CTTCAAGCAGCAGGATTTGTTGTCAACATCGTCGAAGAAGCCAGCGCCGACGTCGCCGAAGGCCTCGTCATCCGAGC
AAACCCAAGCGTTGGATCCGAAATCCGCCAAGGGGCCACCGTCACCATCACCGTGTCCACCGGCCGAGAAATGATCA
ACATCCCAGACGTCTCCGGCATGACACTTGAGGACGCCGCCCGCGCCCTCGAAGACGTTGGTCTCATACTCAACCAA
AACGTTCGGGAAGAAACCTCCGACGACGTCGAATCTGGCCTCGTCATCGACCAAAACCCCGAAGCCGGCCAAGAAGT
AGTCGTGGGTTCCTCTGTATCTCTAACCATGTCTTCAGGCACCGAGAGCATCCGAGTGCCCAACCTCACCGGCATGA
ACTGGTCACAAGCAGAACAAAACCTCATCTCCATGGGCTTTAACCCCACAGCTTCCTACTTAGACAGCAGCGAACCA
GAAGGCGAAGTCCTCTCAGTTTCCAGCCAAGGAACTGAACTACCCAAGGGTTCATCCATCACAGTGGAAGTCTCCAA
CGGCATGCTCATCCAAGCCCCCGATCTCGCCCGCATGTCCACCGAACAGGCCATCAGTGCCCTCCGCGCTGCTGGCT
GGACCGCCCCAGATCAATCCCTGATCGTCGGCGACCCCATCCACACCGCAGCCCTCGTGGATCAAAACAAAATCGGA
TTCCAATCCCCAACCCCTGCAACCCTCTTCCGCAAAGACGCCCAAGTGCAAGTGCGACTCTTCGAATTCGATCTCGC
TGCACTCGTGCAA
>RXA01826-nucleotide sequence C: downstream
TAGCCAACAAGGAAACCGTCAAG

>>RXA01830-amino acid sequence
(1-1353, translated) 451 residues
MLKLKYAVAS DRGLVRGNNE DSAYAGPHLL ALADGMGGHA AGEIASQTMI NHLRALDVDP GDNDMLALVG
MVAGEANAAI AEGIAEDPAR DGMGTTLTAF MFNGRDLAMC HVGDSRGYVL RDDKLVQVTV DDTFVQSLVA
EGKLDPEDVS THPQRSLILK AYTGHPVEPT LEQFPALPGD RLLLCSDGLS DPVTHSTIEE TVRVGTPQDA
STKLVELALR SGGPDNVTVI VADVVEVTEA EAAAEASVPV TAGALNGEQP EDPRPDTAAG RAAAITRRAQ
VIDPAPKISD AGTEDIPTIE EPPEKSSSKL AVLIVALVIL IGVVAAGWWG YSRIDSTFYV AVNDEEAITV
EHGVDYRIFG KDLHSQFQVA CLNEAGTLSL KESCENGTSF KLDDLPASVR GSVAGLPSGS YDEVQAQMQR
LAAQALPVCV NLEVTTGGDR NEPGVNCREV S
>RXA01830-nucleotide sequence A: upstream
ACGGCACTTTTGTCGGTGGTACGCGCATTGATCAGCCTGAGCAGATTGCGGTGGGCACGGATATCCGTATTGGTCGT
ACAGCAGTGAGGCTTGTTCCCTG
>RXA01830-nucleotide sequence B: coding region
ATGTTGAAACTTAAATATGCGGTGGCATCTGACCGAGGGTTAGTGCGCGGGAACAATGAGGATTCCGCTTACGCTGG
CCCGCATTTGTTGGCGCTGGCTGATGGTATGGGCGGCCATGCTGCTGGTGAGATCGCTTCCCAAACCATGATCAACC
ATTTGCGTGCGCTTGATGTTGATCCTGGTGATAACGATATGTTGGCGCTGGTGGGCATGGTGGCAGGCGAAGCCAAC
GCGGCGATTGCTGAGGGCATCGCCGAAGACCCGGCGCGCGACGGCATGGGCACTACGTTGACGGCGTTCATGTTTAA
CGGGCGTGACCTGGCAATGTGCCACGTCGGCGATAGTCGTGGTTATGTGCTTCGCGACGATAAGTTGGTACAGGTTA
CAGTCGACGATACTTTTGTGCAGTCGTTGGTCGCTGAGGGCAAGCTTGATCCAGAAGATGTTTCAACTCACCCTCAG
CGTTCTTTGATTCTGAAGGCTTACACCGGCCATCCTGTGGAGCCCACTCTGGAGCAATTCCCGGCCTTGCCTGGGGA
TCGTTTGTTGTTGTGCTCTGATGGTCTATCAGATCCGGTTACACACTCCACGATTGAAGAAACAGTGCGTGTAGGCA
CCCCGCAGGATGCGTCCACCAAGTTGGTGGAGTTGGCGCTGCGTTCTGGCGGTCCGGACAATGTGACGGTCATTGTG
GCCGATGTTGTAGAAGTCACCGAGGCGGAAGCAGCAGCGGAAGCATCAGTGCCTGTCACGGCTGGTGCGCTCAATGG
TGAGCAGCCTGAAGATCCGCGGCCTGATACCGCTGCGGGACGCGCTGCGGCGATCACACGGCGAGCTCAAGTGATTG
ATCCGGCACCAAAGATATCTGATGCTGGAACGGAGGATATTCCCACAATTGAGGAGCCACCAGAGAAAAGTTCCAGC
AAACTTGCGGTATTGATCGTAGCCCTGGTCATCCTCATCGGTGTAGTTGCCGCAGGATGGTGGGGCTACTCCCGTAT
TGACAGCACTTTTTACGTCGCGGTCAATGATGAGGAAGCCATCACCGTGGAACACGGTGTGGATTACCGCATCTTTG
GCAAGGATTTACATTCGCAATTCCAGGTGGCGTGCCTGAATGAAGCTGGCACCTTGTCACTCAAGGAATCCTGTGAA
AACGGTACGTCTTTCAAATTGGATGATTTACCGGCATCTGTTCGCGGTAGTGTCGCAGGATTACCGTCTGGGTCGTA
TGACGAGGTCCAGGCGCAAATGCAACGGCTGGCTGCTCAAGCTTTGCCAGTGTGCGTGAACTTAGAAGTAACAACCG
GTGGCGATAGAAACGAACCCGGAGTCAATTGTAGGGAGGTCTCA
>RXA01830-nucleotide sequence C: downstream
TGAACACGCTTGAACGATTAAAG

>>RXA00786-amino acid sequence
(1-879, translated) 293 residues
TGKKILLAAP RGYCAGVDRA VETVERALEE YGAPIYVRKE IVHNRYVVDT LAEKGAIFVN EASEAPEGAN
MVFSAHGVSP MVHEEAAAKN IKAIDAACPL VTKVHKEVQR FDKQGFHILF IGHEGHEEVE GTMGHSVEKT
HLVDGVAGIA TLPEFLNDEP NLIWLSQTTL SVDETMEIVR ELKVKFPQLQ DPPSDDICYA TQNRQVAVKA
IAERCELMIV VGSRNSSNSV RLVEVAKQNG ADNAYLVDYA REIDPAWFEG VETIGISSGA SVPEILVQGV
IERLAEFGYD DVE
>RXA00786-nucleotide sequence B: coding region
ACCGGAAAGAAGATCCTGCTTGCAGCCCCTCGCGGATACTGTGCCGGCGTAGACCGTGCAGTGGAAACCGTCGAGCG
CGCGCTCGAGGAATACGGCGCCCCAATTTATGTCCGTAAAGAAATCGTGCACAACCGTTACGTTGTGGACACCCTGG
CAGAAAAGGGCGCGATTTTTGTCAACGAAGCATCTGAAGCACCAGAAGGTGCCAACATGGTGTTCTCTGCACACGGC
GTGAGCCCAATGGTCCACGAAGAAGCTGCAGCTAAAAACATCAAGGCTATTGACGCGGCCTGCCCGCTGGTCACCAA
AGTGCACAAGGAAGTCCAGCGCTTTGATAAGCAGGGATTCCACATTCTCTTCATCGGTCACGAAGGCCATGAAGAAG
TAGAGGGCACCATGGGTCATTCCGTTGAGAAAACCCACCTGGTTGACGGCGTTGCTGGCATTGCCACCCTGCCTGAA
TTCTTAAACGATGAACCAAACCTGATCTGGCTGTCTCAGACCACGCTTTCTGTGGACGAGACCATGGAGATCGTCCG
CGAGCTGAAGGTGAAGTTCCCTCAGCTGCAGGATCCACCGTCAGATGATATTTGCTACGCCACGCAGAACCGCCAGG
TTGCCGTCAAGGCTATCGCTGAGCGCTGCGAGCTGATGATTGTGGTCGGTTCCCGCAACTCCTCCAACTCGGTTCGT
CTGGTTGAGGTCGCTAAGCAAAACGGTGCCGATAACGCCTACCTGGTGGATTACGCCCGCGAAATCGACCCAGCATG
GTTCGAAGGCGTAGAGACCATCGGTATCTCCTCCGGCGCTTCCGTGCCTGAGATCCTCGTCCAGGGCGTCATTGAGC
GCCTGGCTGAGTTCGGCTACGACGACGTCGAG

>>RXA00319-amino acid sequence
(1-426, translated) 142 residues
MSEVIAKAKA EEAGLEDNVI FSSCGMGNWH VGQPADKRAL AELKSAGYNG DTHRAAQLGP EHMRADLFVA
LDSGHAGELA ATGVPNDKIR LMRSFDPESN PTDDVADPYY GTSQDFVLTR ENIEDAMPGL LEWVRDHIRT
DS
>RXA00319-nucleotide sequence A: upstream
ACCGCAGAAGAATTAGAAAGGATCATCCATGACTGGGCCTAAAACTTCGCTACCTGTGGAAATTGTTTTCGTATGCA
CCGGAAACATTTGCCGATCCCCC
>RXA00319-nucleotide sequence B: coding region
ATGTCGGAAGTCATCGCGAAGGCAAAAGCGGAAGAAGCTGGCTTGGAAGACAACGTCATTTTCTCCTCCTGTGGCAT
GGGCAATTGGCACGTTGGCCAACCTGCTGACAAGCGAGCTCTCGCGGAACTGAAATCAGCCGGTTACAACGGCGACA
CCCACCGCGCAGCACAACTTGGTCCCGAGCACATGCGCGCAGATCTCTTCGTCGCGCTAGATTCCGGCCACGCCGGT
GAGCTCGCCGCAACGGGTGTTCCCAACGACAAAATCCGCCTCATGCGTTCCTTCGACCCAGAGTCCAACCCCACCGA
CGATGTCGCAGACCCTTACTACGGCACATCCCAGGATTTCGTGCTCACCCGTGAAAACATCGAAGATGCTATGCCGG
GCCTTTTGGAGTGGGTCAGAGATCACATCCGCACTGATTCT
>RXA00319-nucleotide sequence C: downstream
TAGGTCTTTGAGCTAAAAAGTCC

Claims

1. An isolated nucleic acid molecule from *Corynebacterium glutamicum* encoding a metabolic pathway regulatory protein, or a portion thereof.

2. The isolated nucleic acid molecule of claim 1, wherein said metabolic pathway regulatory protein is selected from the group consisting of proteins involved in the regulation of metabolism of organic acids, proteinogenic and nonproteinogenic amino acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated and unsaturated fatty acids, diols, carbohydrates, aromatic compounds, vitamins, cofactors, and enzymes.

3. An isolated *Corynebacterium glutamicum* nucleic acid molecule selected from the group consisting of those sequences set forth in Appendix A, or a portion thereof.

4. An isolated nucleic acid molecule which encodes a polypeptide sequence selected from the group consisting of those sequences set forth in Appendix B.

5. An isolated nucleic acid molecule which encodes a naturally occurring allelic variant of a polypeptide selected from the group of amino acid sequences consisting of those sequences set forth in Appendix B.

6. An isolated nucleic acid molecule comprising a nucleotide sequence which is at least 50% homologous to a nucleotide sequence selected from the group consisting of those sequences set forth in Appendix A, or a portion thereof.

7. An isolated nucleic acid molecule comprising a fragment of at least 15 nucleotides of a nucleic acid comprising a nucleotide sequence selected from the group consisting of those sequences set forth in Appendix A.

8. An isolated nucleic acid molecule which hybridizes to the nucleic acid molecule of any one of claims 1-7 under stringent conditions.

9. An isolated nucleic acid molecule comprising the nucleic acid molecule of any one of claims 1-8 or a portion thereof and a nucleotide sequence encoding a heterologous polypeptide.

10. A vector comprising the nucleic acid molecule of any one of claims 1-9.

11. The vector of claim 10, which is an expression vector.

12. A host cell transfected with the expression vector of claim 11.

13. The host cell of claim 12, wherein said cell is a microorganism.

14. The host cell of claim 13, wherein said cell belongs to the genus *Corynebacterium* or *Brevibacterium*.

15. The host cell of claim 12, wherein the expression of said nucleic acid molecule results in the modulation in production of a fine chemical from said cell.

16. The host cell of claim 15, wherein said fine chemical is selected from the group consisting of: organic acids, proteinogenic and nonproteinogenic amino acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated and unsaturated fatty acids, diols, carbohydrates, aromatic compounds, vitamins, cofactors, and enzymes.

17. A method of producing a polypeptide comprising culturing the host cell of claim 12 in an appropriate culture medium to, thereby, produce the polypeptide.

18. An isolated metabolic pathway regulatory polypeptide from *Corynebacterium glutamicum*, or a portion thereof.

19. The protein of claim 18, wherein said polypeptide is selected from the group of metabolic pathway proteins which participate in the regulation of metabolism of organic acids, proteinogenic and nonproteinogenic amino acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated and unsaturated fatty acids, diols, carbohydrates, aromatic compounds, vitamins, cofactors, and enzymes.

20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of those sequences set forth in Appendix B.

21. An isolated polypeptide comprising a naturally occurring allelic variant of a polypeptide comprising an amino acid sequence selected from the group consisting of those sequences set forth in Appendix B, or a portion thereof.

22. The isolated polypeptide of any of claims 18-21, further comprising heterologous amino acid sequences.

23. An isolated polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence which is at least 50% homologous to a nucleic acid selected from the group consisting of those sequences set forth in Appendix A.

24. An isolated polypeptide comprising an amino acid sequence which is at least 50% homologous to an amino acid sequence selected from the group consisting of those sequences set forth in Appendix B.

25. A method for producing a fine chemical, comprising culturing a cell containing a vector of claim 12 such that the fine chemical is produced.

26. The method of claim 25, wherein said method further comprises the step of recovering the fine chemical from said culture.

27. The method of claim 25, wherein said method further comprises the step of transfecting said cell with the vector of claim 11 to result in a cell containing said vector.

28. The method of claim 25, wherein said cell belongs to the genus *Corynebacterium* or *Brevibacterium*.

29. The method of claim 25, wherein said cell is selected from the group consisting of: *Corynebacterium glutamicum*, *Corynebacterium herculis*, *Corynebacterium lilium*, *Corynebacterium acetoacidophilum*, *Corynebacterium acetoglutamicum*, *Corynebacterium acetophilum*, *Corynebacterium ammoniagenes*, *Corynebacterium fujiokense*, *Corynebacterium nitrilophilus*, *Brevibacterium ammoniagenes*, *Brevibacterium butanicum*, *Brevibacterium divaricatum*, *Brevibacterium flavum*, *Brevibacterium healii*, *Brevibacterium ketoglutamicum*, *Brevibacterium ketosoreductum*, *Brevibacterium lactofermentum*, *Brevibacterium linens*, *Brevibacterium paraffinolyticum*, and those strains set forth in Table 3.

30. The method of claim 25, wherein expression of the nucleic acid molecule from said vector results in modulation of production of said fine chemical.

31. The method of claim 25, wherein said fine chemical is selected from the group consisting of: organic acids, proteinogenic and nonproteinogenic amino acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated and unsaturated fatty acids, diols, carbohydrates, aromatic compounds, vitamins, cofactors, and enzymes.

32. The method of claim 25, wherein said fine chemical is an amino acid.

33. The method of claim 32, wherein said amino acid is drawn from the group consisting of: lysine, glutamate, glutamine, alanine, aspartate, glycine, serine, threonine, methionine, cysteine, valine, leucine, isoleucine, arginine, proline, histidine, tyrosine, phenylalanine, and tryptophan.

34. A method for producing a fine chemical, comprising culturing a cell whose genomic DNA has been altered by the inclusion of a nucleic acid molecule of any one of claims 1-9.

CORYNEBACTERIUM GLUTAMICUM GENES ENCODING REGULATORY PROTEINS

Abstract of the Disclosure

Isolated nucleic acid molecules, designated MR nucleic acid molecules, which encode novel MR proteins from *Corynebacterium glutamicum* are described. The invention also provides antisense nucleic acid molecules, recombinant expression vectors containing MR nucleic acid molecules, and host cells into which the expression vectors have been introduced. The invention still further provides isolated MR proteins, mutated MR proteins, fusion proteins, antigenic peptides and methods for the improvement of production of a desired compound from *C glutamicum* based on genetic engineering of MR genes in this organism.